International Initiative for Impact Evaluation

Systematic Review of the Effectiveness of Contract Farming for Income and Food Security of Smallholder Farmers in Low- and Middle Income Countries

3ie Systematic Review – SR6.1088
Protocol

18 December 2015

Review team:
dr. Giel Ton, LEI Wageningen UR – Lead Reviewer
dr. Wytse Vellema, Ghent University – Reviewer
prof. dr. Marijke D'Haese, Ghent University – Reviewer
dr. Sam Desiere, Ghent University – Reviewer
\Sophia Weituschat – Review assistant
Ans Brouwer/ Hans van Hoeven – Information Specialists

Contact information:
Giel Ton
LEI Wageningen UR
P.O. Box 29703
The Hague, 2502 LS, The Netherlands
giel.ton@wur.nl

Wytse Vellema
Ghent University
Department of Agricultural Economics
Coupure Links 653
Gent, 9000, Belgium
wytse.vellema@ugent.be

Funding:
Research discussed in this publication has been funded by the International Initiative for Impact Evaluation, Inc. (3ie), provided under Systematic Review – 6 Window

1. Background

Motivation for the review

Contract farming is a commercial relationship between a firm and a group of farmers. It is a business model in which production is bought in advance by a firm in exchange for certain services and other benefits. Although principally a commercial initiative, contract farming is considered to be a way to overcome the challenges that small farmers face when linking to remunerative markets. It assists farmers in connecting to output markets and often provides inputs, credit, or agricultural extension (Eaton and Shepherd, 2001; World Bank, 2007; Da Silva and Ranking, 2013). These services can be not only provided by private firms, but can also come from, or be facilitated by, multi-actor partnerships between companies, governments and NGOs (Prowse, 2012). Estimates of the incidence of contract farming in developing countries are unreliable and differ markedly between countries; they are generally below 10% of total area under production (Minot and Ronchi, 2015). A rapidly growing number of firms – at least in modern market channels – are increasingly relying on contract procurement (Da Silva and Ranking, 2013).

Companies may offer a contract only to those farmers who comply with some minimum requirements (for example land ownership, irrigated lands, minimal plot sizes, etc.) which may preclude smaller farms from benefitting directly. Even if these arrangements are beneficial for smallholders directly or indirectly through spill-over effects, there will likely be heterogeneity in impacts, with certain farmers benefitting while others do not or even lose out. It is clear that contracts will not be randomly distributed within a farming community, and contracting farmers will always have special characteristics, or so-called firm-selection and self-selection biases (Minot and Ronchi, 2015; Barrett et al., 2012). Studies that infer quantitative effects of contract farming on income and food security will need a proper research design to control for these biases.

Public policies and institutions influence the market strategies and forms of transaction used between farmers and firms. Contracting firms are almost always relatively large processors, exporters, or supermarket chains. Rarely do small-scale traders or even wholesalers offer farmers pre-planting contracts. This is not surprising, given the large fixed costs associated with contracting (Minot and Ronchi, 2015). Contract farming is induced by a firm's need to source products with specific qualities and in sufficient quantities (Minot and Ronchi, 2015) and the presence of appropriate geographical and political-economic conditions and an enabling business environment (Jia and Bijman, 2013). Relevant geographical conditions are road infrastructure, access to water, soil types, climate, etc. Relevant political-economic conditions are land-rights policies, market regulation, trade policies and the risk of socio-economic shocks. Whether a firm chooses to start offering a contract farming arrangement is also highly influenced by the local business environment, such as financial services, conflict resolution systems, investment subsidies, business development services, brokering services, and farmer organisations. Many of the above-mentioned conditions are influenced by political decision-making. Therefore, policy makers can enable or constrain the opportunities for contract farming, influencing its attractiveness to firms and farmers as a way of organising transactions and embedded services. Since 2007, coinciding with the investor rush for land in sub-Saharan Africa, international development agencies have increasingly presented contract farming as an alternative or complementary development opportunity for smallholder inclusion (Lindholm, 2014).

Contract farming is considered by most authors to be a positive development for the inclusion of farmers in markets (Minot, 1986; Eaton and Shepherd, 2001), food security (Bellemare and Novak, 2015), and global poverty reduction (Setboonsarng and Leung, 2014). Yet, there is serious concern whether smallholders are able to benefit from these arrangements, because the relative size of buyers most likely results in an unequal power relationship, which influences the terms of the arrangements (Sivramkrishna and Jyotishi, 2008; von Hagen and Alvarez, 2011).

In the earlier literature on contract farming, the issue of power imbalances was especially prominent (Little and Watts, 1994), and the discussion was rather polarised between proponents and critics of contract farming (Oya, 2012). Recent literature, however, has paid more attention to diversity within contract farming arrangements, and, for example, acknowledges the complex interactions involved in a contract farming scheme and how farmers and firms make use of their constrained room for manoeuvre (Vellema, 2002; Minot and Ronchi, 2015).

Though, contract farming is the outcome of a complex process influenced by many different factors and in which policies and projects shape the enabling environment in which a firm and a farmer agree on a contract, policy makers, international organisations propose (Lindholm, 2014), and many scholars study (Minot and Ronchi, 2015; Da Silva and Ranking, 2013), contract farming as a way to improve the inclusion of farmer in markets, and consider it a development strategy. Although it is not a well-defined, standard 'treatment', we think that a comparison of these effectiveness studies on various empirical instances of contract farming will give an overview of the evidence base behind support to contract farming as a development strategy and can derive useful policy recommendations that have wider application.

Many studies in the past compared the incomes of participating and non-participating farmers, or incomes before and after the signing of the contract, and have a high risk of selection bias: e.g., rich farmers may participate more than poor farmers, etc. This results in unreliable effect estimates that are more a reflection of selection bias or confounding factors (e.g., weather or prices) than of the contract farming arrangement itself. The last decade shows a rapid increase of studies on contract farming that assess the effects of contract farming using quasi-experimental research designs, which provide much more reliable estimates. As can be expected, these studies report mixed effects (Minot and Ronchi, 2015).

These impact evaluations assess the effects of specific instances of contract farming on smallholders. Therefore, a systematic review of this rapidly growing body of evidence is timely, and may help to distil generalised inferences from these specific instances. To do so, our review follows a two-stage process. First, it identifies and discusses the results of impact studies that applied a research design to reduce selection bias and assess the counterfactual situation (What would have happened to the smallholder farmers had there been no contract farming arrangement?). The review compares the inferences in these studies and assesses the strength of the evidence behind them, as well as the rigour of the methods used. Secondly, it places these contractual arrangements in context. The review maps the relevant contextual conditions and elements of the enabling environment for each of the empirical instances of contract farming covered by these studies, searching for missing information through a targeted information search around each of these instances of contract farming, and making a case-based comparative analysis in order to distill recommendations for policy makers and practitioners.

Description of Contract Farming

Contractual arrangements in agriculture can be extremely diverse, with varying embedded services, credit arrangements, payment systems and price-setting mechanisms. Contract farming is generally defined very broadly as 'agricultural production carried out according to an agreement between a buyer and farmers, which establishes conditions for the production and marketing of a farm product or products' (http://www.fao.org/ag/ags/contract-farming/faq). We focus on a more strict definition of contract farming, partly based on Prowse (2012), where there is a direct sales agreement between a firm and a farmer before production begins, and the farmer gets seeds, inputs or credit to pay his or her production costs.

In this review, contract farming is defined as, *"a contractual arrangement for a fixed term between a farmer and a firm, agreed verbally or in writing, before production begins, which provides material or financial resources to the farmer and specifies one or more product or process requirements, for agricultural production on land owned or controlled by the farmer, which gives the firm legal title to (most of) the crop"*.

There are many different types of contracts, going from full resource provisioning contracts with detailed production and marketing conditions to a mere verbal agreement to buy whatever quantity is produced at the going market price. Therefore, for each empirical instance of contract farming in the studies selected for synthesis, we will map the differences in embedded services provided and contractual conditions, such as payment systems, quality control, dispute resolution, etc. included in each of the contracts. We note these contextual conditions in order to find predictors of effectiveness for smallholders, and to distil good practices that can be used to fine-tune the contracts with the aim of increasing smallholder income and food security.

Theory of Change for contract farming

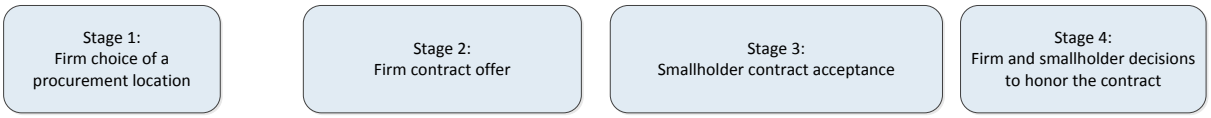
This review will synthesise the documented positive, negative, intended, and unintended outcomes of the contract farming arrangements. The intended effects of the contractual arrangements for the contracting farmers are to be found in the outcomes directly related to the contracted production, such as improved agricultural practices, improved yields and improved crop revenue. Non-contracted farmers may experience spill-over effects (e.g., information and technology obtained from neighbours). Contract farming might also have negative outcomes, including exposure to financial risk, health and food safety. Most effects can be expected only when the contract is continued for several years by firm and farmers. However, even if one of the parties decides to break or discontinue the contract after a short time, some changes in agricultural practices may have resulted in positive or negative intermediate or ultimate outcomes.

As in Ton et al.'s 2015 review of the impact research on innovation grants (Ton et al., 2015; Ton et al., 2013), we will outline the range of proxy-indicators, outcome areas and effects that have been used to assess the impact of contract farming on the livelihood strategies of smallholder farmers.

Outcome areas

Following Barrett et al (2012), we identify four stages of decision-making by contracting parties concerned that are required to lead to outcomes for smallholders.

Figure 1. Process of establishing a contract farming arrangement

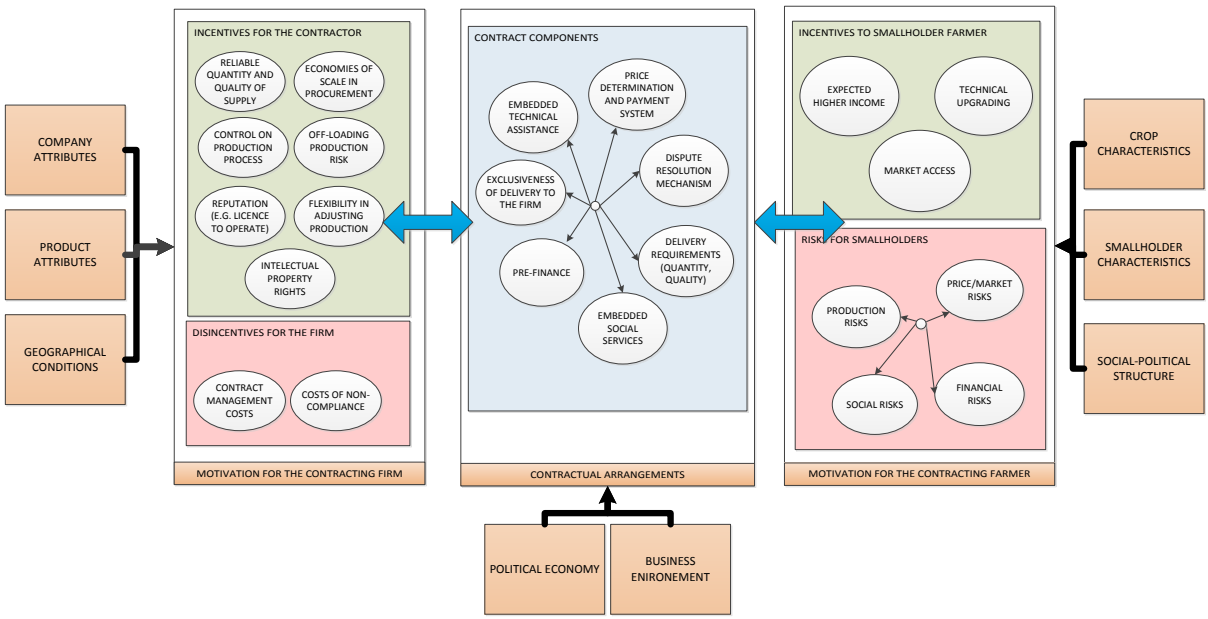


In stage 1, a firm chooses their procurement location; next, in stage 2, the firm proposes a contract to a (specific group) of farmers, which these farmers may accept. The firm may deliberately restrict the contract to those farmers complying with specific characteristics. Stage 3 is the acceptance of the contract details by the farmer(s). The firm and smallholder then decide to honour the contract for its duration, which we consider to be stage 4. A farmer organisation, (local) government, financial institution and/or NGO may play a role in brokering this agreement. Subsequently, farmers accept the proposal, and agree to be part of the contractual arrangement. After each agricultural cycle or contract period, the parties may decide to either stop or renew their arrangements. The latter is dependent on the fixed investments made by each of the contracting parties, and the alternative use of these investments outside the contract arrangement. For example, discontinuation of a contract with a firm after having planted perennial crops (e.g., for biofuels) will imply higher costs than when stopping a specific (variety of) vegetable production.

Immediate outcomes of a contract farming arrangement are, thus, related to the uptake and renewal of contracts. The length of time that the contractual arrangement has been active will be taken into account in the narrative discussion of findings in each study. The continuation of a contractual arrangement is an indicator of an overall satisfaction of the firm and the farmers with their contractual arrangement. Discontinuation of a contract (drop-outs) could signal that, at least in the short run, the firm or the farmer consider it suboptimal and not in their interest to renew. There may also be side-selling, where farmers sell to another buyer than the one with whom they are contracted. Figure 2 illustrates the incentive structure for firms and farmers to decide to enter and renew a contractual arrangement. This negotiation process implies that, in time, the content of the contract and the characteristics of the contracting parties will change.

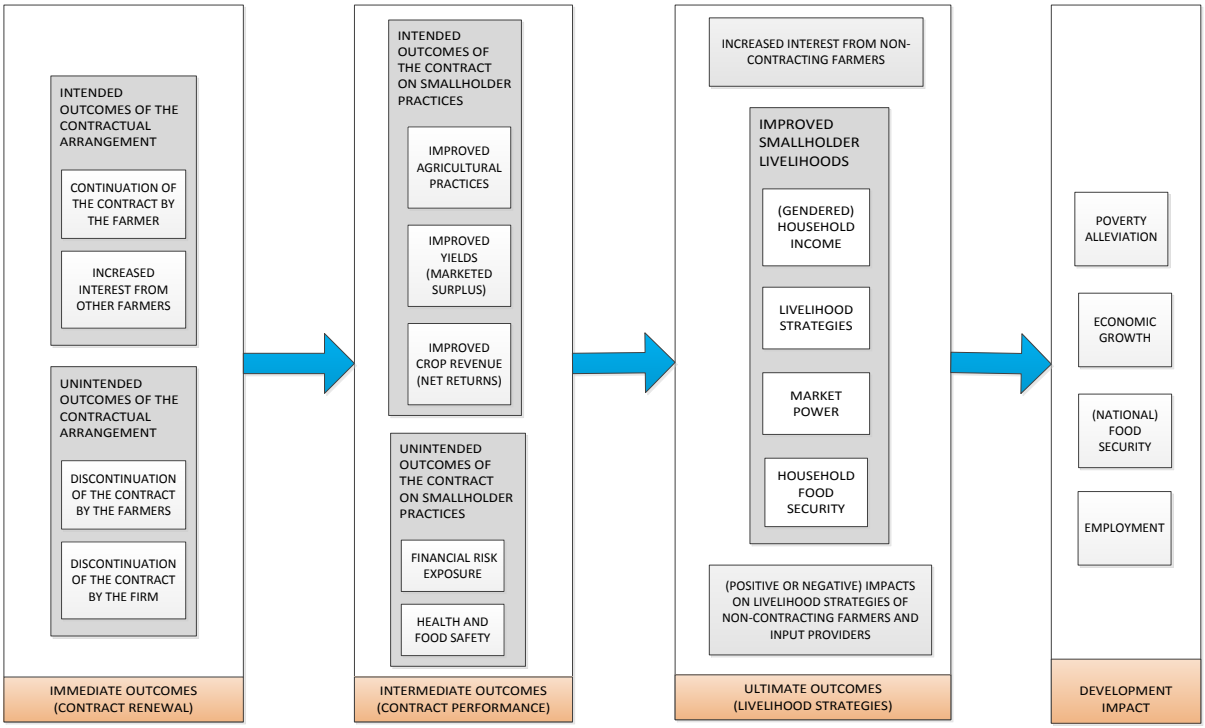
Effects on smallholders will be greater for contract farming arrangements that are in place for a longer time, because investments in productive assets and knowledge take time to bear fruit. It is likely that unsuccessful farmers will have left the arrangement. Research by Narayanan (2013) in India suggests that farmers may move in and out of contracts quite regularly, thus indicating a potential difference between observed short-term dynamics and a more structural mid-term effect on the (local) institutional environment. In the long run, contractual arrangements between firms and farmers may become a common phenomenon, and one of the available business options for production/marketing.

Figure 2. Contractual arrangements as a result of a negotiation process and influencing factors.



Intermediate outcomes are indicators of contract performance and productivity. Input-providing contracts are believed to be most beneficial for the poorest farmers, as they lack the financial capacity to invest in these inputs themselves and they do not have access to finances that would enable them to buy these inputs on credit. In this way, contract farming becomes a form of 'inclusive business' (Vorley and Proctor, 2008). Ultimately, the increased value of agricultural production under contract should increase household income, or could increase household food security by providing more stable or less uncertain incomes. Finally, improved productivity in the contracted crop could spread to other crops, resulting in an additional increase in income, or, if the crops are for home consumption, improved household food security (Minten et al., 2009). Effects on intermediate outcomes can be visible over a shorter time-span than changes in ultimate outcomes, such as household income and food security. It is nevertheless more challenging to compare the immediate and intermediate outcomes between different contract farming instances, because the indicators used to measure these outcomes tend to be context- and product specific and rely on different constructs and ways of measurement. Our review will systematically map the proxy-indicators used by each author and try to apply meta-analysis on various impact areas (grouping similar proxy-indicators) in the chain of effect- immediate, intermediate and ultimate outcomes – indicating if and why the contexts, treatment and outcomes are deemed sufficiently comparable. We set out the following intended intermediate outcome areas with our coding tool (see Figure 3): improved agricultural practices, improved yields, and improved crop revenue. As unintended intermediate outcomes, we will map information on financial risk exposure, health and food safety. Our meta-analysis will be limited to the intended outcomes, as these are most commonly measured in the studies on effectiveness.

Figure 3. Outcome areas that result from the contractual arrangement



The ultimate outcomes of contract farming arrangements are income and food security of smallholder farmers. The causal link between changes in these ultimate outcomes and the contract farming arrangements is not always clearcut, especially when contract farming covers only a small part of a farmer's agricultural activities. For smaller contract farming schemes, net-effects on ultimate outcomes might be difficult to detect with quasi-experimental research designs, due to the (likely) lower effect size (Ton et al., 2014b). We map the following ultimate outcome areas with our coding tool in order to conduct the meta-analysis (see Figure 2): (intra-) household income, market power, and household food security. Other outcomes on livelihood strategies of participating and non-participation farmers are discussed in the narrative synthesis.

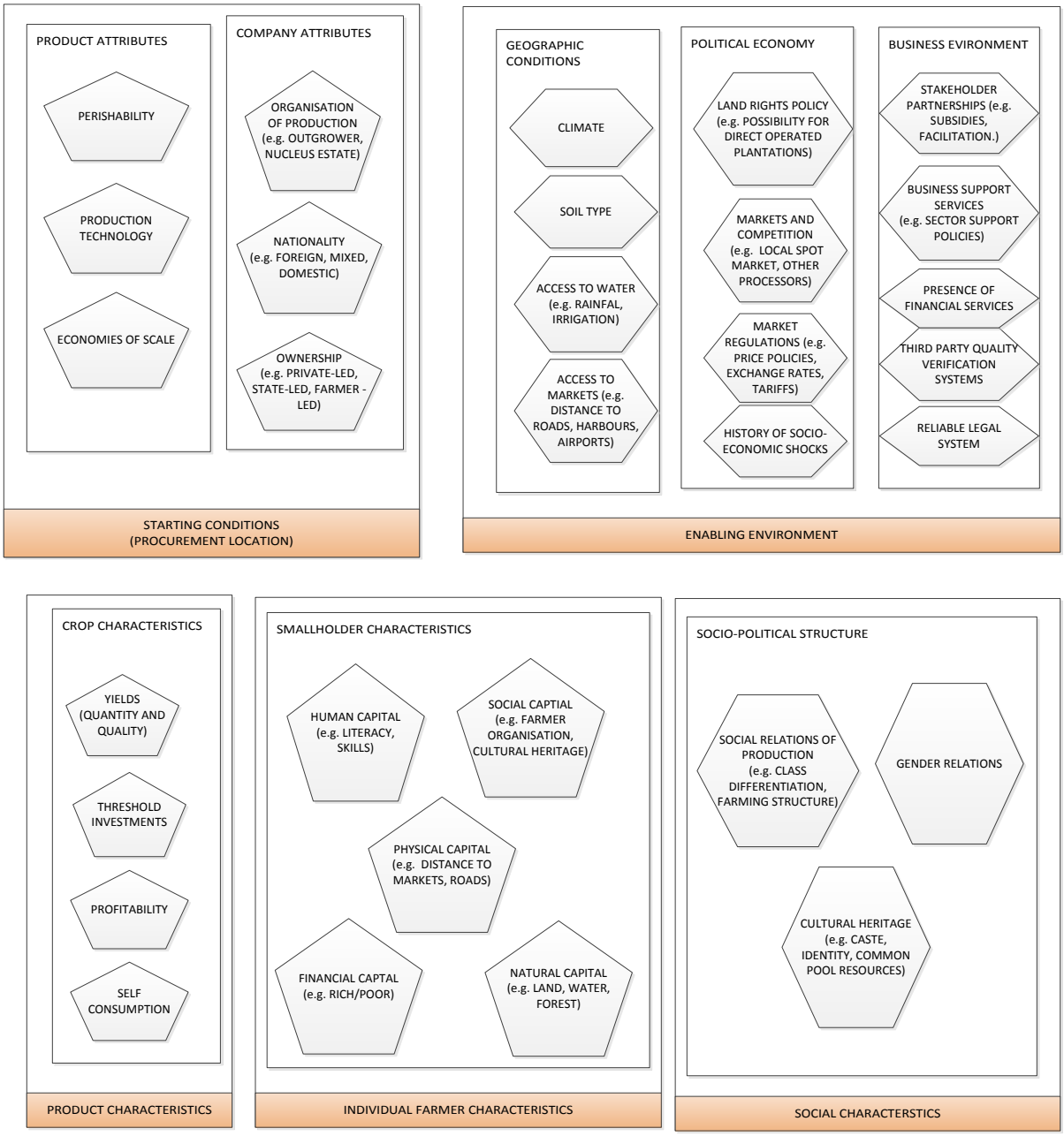
Whenever the studies give empirical evidence about development impact, e.g. poverty alleviation, economic growth, national food security and employment, we will distil this information, and use it in a narrative discussion about each empirical instance. It is unlikely that there are studies that quantify effects at this outcome level; most study will narratively explain if and how the intermediate and ultimate outcomes contributed to these broader outcome areas. There are also expected outcomes for traders, suppliers and non-beneficiary farmers, on which available evidence will be collected (see below).

Barriers and enablers

We will explore plausible enablers and barriers of effectiveness in the instances of contract farming for which good-quality evidence on effects is available. Figure 4 includes the conditions and characteristics that we consider potential enablers or barriers: the policy context, enabling environment, and the characteristics of the firms, the farmers and the crops. This list is not exclusive. For each empirical instance, we will search for information on these influencing factors. Most of this

information will be qualitative in nature. This part of the study is geared to making policy recommendations about potential 'recipes' of enabling conditions that can be expected to improve the positive effects of contract farming in income and food security of smallholders, and reduce the negative effects. These factors are further refined in the coding tool provided in the methods section.

Figure 4. Factors incorporated in the coding tool to explore for predictors of effectiveness



Why it is Important to do This Review

The body of literature on contract farming is growing rapidly. Moreover, a preliminary scan of the literature revealed that there is a great heterogeneity in contract farming, with differences in contracts, farmers, products, buyers, and institutional environments. In spite of this heterogeneity, a systematic review of these studies is important, in order to extract some generalised inferences from the accumulated knowledge about the effects of contract farming, intended and unintended, direct and indirect. The review will synthesise the research which has been done, identifying commonalities and differences, generalizable and context-specific findings, and exploring the evidence for lessons learned that may improve current and future contract farming arrangements.

It is important to improve the context for contract farming arrangements in order to generate positive effects for smallholder farmers, in the light of the existing body of evidence. The Population Division of the UN department of Economic and Social Affairs expects the world population to reach 9.6 billion by 2050, of which 6.3 billion are expected to live in cities (United Nations Population Division, 2014).

The agricultural production required to feed this growing and mostly urban population needs to come from more intensive use of land already used for agriculture, as there is little unused land left in the world. Farms smaller than five hectares occupy almost three-quarters of farmland in low-income countries and almost two-thirds in the lower-middle income countries (FAO IFAD WFP, 2014). Most of these farmers are poor (World Bank, 2007), which makes research to better understand the effectiveness of contract farming and the ways that it can improve their market power, income, and food security a potentially relevant instrument for poverty alleviation. We are not aware of any completed systematic reviews on the effectiveness of contract farming to overcome these challenges.[1]

The review is timely. Policy makers are increasingly interested in working with the private sector in development initiatives, using co-investment or developing joint projects. The involvement of the private sector in the international development agenda is growing rapidly. A good example of this is the World Business Council for Sustainable Development, but there are many similar initiatives. Therefore, this review not only aims to map the evidence about the effects that contract farming has had on income and food security of smallholder farmers, according to studies with a strong research design, but it aims also to show what seems to be working, when, and under which conditions. It is also a unique opportunity to identify knowledge gaps to direct further research on the topic.

[1] There is an on-going Campbell review on the effectiveness of agricultural certification schemes (Oya et al., 2015).

2. Objectives

Our systematic review will use the following questions to guide the review and synthesis process. These questions are divided into two sets.

Question 1: What is known about the size of effects of contract farming on income and food security of smallholder farmers in low- and middle income countries?

1.1 What are the effects of the contract farming arrangement on ultimate outcomes such as household income and food security?
1.2 What are the effects of the contract farming arrangement on intermediate outcomes such as agricultural practices, yields and/or net-returns derived from producing the contracted crop?
1.3 What are the effects on intermediate and ultimate outcomes for non-participating neighbouring farmers living in the same communities as the contract farmers?
1.4 What are the drop-out rates and side-selling associated with the contract farming arrangement (immediate outcomes)?

Question 2: Under what enabling or limiting conditions are contract farming arrangements effective for improving income and food security of smallholders?

2.1. What are the enabling and disabling conditions that explain the effectiveness and sustainability of the contract farming arrangements for smallholder famers?
2.2. Are there configurations of conditions that may enable success or failure of contract farming arrangements?

3. Methods

3.1 Criteria for considering studies for this review

In the review we will use several criteria to select studies useful for answering the research questions. As a first step, we aim to identify studies with an experimental or quasi-experimental design that assess the effects of the contract farming on income and food security of smallholder farmers in developing countries. Once these empirical instances have been selected, in a second step we will search for additional information (qualitative studies, background information) to reflect on the conditions and dynamics in the empirical instances associated with these studies, to identify enablers and barriers to the effectiveness of these contract farming arrangements. The quality (rigour) of the methods to support conclusions in these information sources will be assessed using the critical appraisal tool used in the systematic review on farmer field schools (Waddington et al., 2014).

The criteria to select the effectiveness studies in the first step are described below following the PICOS format (Participants, Intervention, Comparisons, Outcomes, Study design).

Types of Participants

The review will focus on the effectiveness of contractual arrangements for smallholders farmers. We will also include studies that assess effects on other stakeholder groups than farmers, whenever these concern agricultural land in areas where also smallholder farmers are located. As the evidence base,

we will include empirical studies that cover contractual arrangements in low-income economies, lower-middle income economies and upper-middle income economies, according to the World Bank, at the time of data-collection. To sketch the range of dimensions relevant for the definition of smallholder farms, we will report the definition used by the authors of the primary studies, and discuss these according to four criteria proposed by Stewart et al. (2014):

1. Limited size of farm (as compared to other farms in the sector)
2. Mostly dependent on family labour
3. Subsistence farming or mix of subsistence and market-oriented farming
4. Reportedly limited resources in terms of land, technical and technological support, and/or capital for maintenance and investment.

Types of Interventions

The interventions studied cover a specific type of contractual arrangement in agriculture, which are agreed for fixed term between a farmer and a firm, verbally or in writing, before production begins for agricultural production on land owned or controlled by the farmer. Such agreements provide material or financial resources to the farmer and specify one or more product or process requirements, and give the firm legal title to the crop. While the intervention is usually initiated between firms and smallholder farmers, we will also include relevant public policies, programmes and projects which aim to promote such practices or improve on the terms and conditions. Following this definition, in our screening procedure, we explicitly exclude:

- Contractual arrangements that have no service-providing clauses; for example, forward sales and price hedging, which are too-common practices. 'Contract farming' as usually referred to by FAO, UNCTAD and World Bank, and as defined by this review, implies that there needs to be some services for technological upgrading of production (Lindholm, 2014).

- Studies on contractual arrangements that only concern marketing (e.g., collective marketing, marketing boards, preferred suppliers, etc.), except when accompanied by pre-harvest service delivery from the procuring firm directly to the contracted smallholders.

- Traditional sharecropping arrangements in which a smallholder works his land on a sharecropping basis with a farmer who provides the necessary inputs in exchange for a part of the harvest. They will be included only if a firm is involved, and the contract is offered to multiple farmers under the same conditions.

- Product certification schemes (e.g., Utz, Fair Trade, Rainforest Alliance) that do not have a non–transferable fixed term forward sale contract with a specific firm; these studies are covered by a separate systematic review that is ongoing (Oya, 2015)

- Hybrid situations, such as a contract between a cooperative and a fixed buyer, and subsequent services delivery of these organisations to the members will only be included if the decision to enter in such a contractual arrangement with the firm is optional for individual farmers/members. This excludes most instances of Fair Trade exports, in which it is the cooperative that sells to a firm, not the individual farmer.

- Contractual arrangements that are not within the realm of 'agriculture', such as forest exploitation and marine fishery.

Types of Comparisons

Contract farming is not a single homogenous intervention but more accurately conceptualised as a package of various services provided to a farmer in exchange for a pre-planting sales agreement with a buyer: higher or more stable prices, a guaranteed market, inputs or financing on credit, agricultural extension, access to improved inputs, certification, and transport. If the contract contains such benefits, it needs to be clear which of these services are also available to the comparison group not having a pre-planting sales agreement. The actual interventions considered in the synthesis will be the differences in service packages made available to farmers in exchange for the pre-planting sales agreements with the buyers.

Comparison-group farmers will usually be located in the same region as intervention farmers. In instances with no geographical separation between intervention and control groups, studies should usually have an analytical design that enables control for spill-over effects, or they will be down-graded according to risk of bias criteria.

The treatment and comparison groups in the studies need to be similar, and to include at least a subgroup of smallholder farmers. Where necessary, the meta-analysis will consider effects on smallholders only, using a disaggregated analysis of the effects on this subgroup.

Types of Outcomes

The meta-analysis of studies of effects will focus on the intermediate and ultimate outcomes of the contractual arrangements (see Chapter 1 on the Theory of Change). We anticipate a wide diversity of proxy-indicators (e.g., survey questions) used by different authors to measure similar outcome areas of contracts. We will focus on the *agricultural practices*, *yields* and/or *net returns derived from producing the contracted crop*. These intermediate outcomes are the most important justification for farmers to enter into a contract farming arrangement.

As ultimate outcomes, we will focus on *households income, market power* and *food security*. We list the proxy-indicators used in each study to register changes in these outcome areas. Income may be measured directly (income, expenditure) or indirectly (poverty, asset, or income indicators). Food security may also be measured by a wide range of indicators (Heady and Ecker, 2013; Bellemare and Novak, 2015). Potential intra-household heterogeneity of effects will be taken into account. These ultimate outcomes are the most important justification for the policies and investments that aim to facilitate and promote contract farming for poverty reduction.

Contracts may also have unintended outcomes; outcomes that change because of the contractual arrangement but that were not the reason for starting the arrangement in the first place. Unintended outcomes include both positive and negative impacts. Positive *unintended outcomes* are spill-over effects to non-participating farmers, either through transfer of knowledge or employment generation, or unintended outcomes in the household as a result of the contractual arrangement, such as educational opportunities, health care, or increased visibility by and access to other development

support. Potential negative *unintended outcomes* include financial risk exposure (indebtedness), health (agrochemicals), reduced on-farm biodiversity (mono-cropping). Positive or negative changes to intra-household gender relations may also occur.

Study designs

Most studies evaluating the impact of contract farming on smallholder producers analyse a single contractual arrangement (case study). Such evaluation studies may assess whether or not a specific arrangement was successful. However, because of the case study context, they are limited in their ability to make inferences about the effectiveness of various specifications of the contractual arrangement and to point to combination of factors that may be conducive in different contexts. We expect to make a comparative study of empirical instances of contract farming for which good-quality effectiveness studies have been done. In the review, we hope to identify, through a case-based comparative analysis, several contextual factors and contract characteristics that seem to be key ingredients for an enabling institutional and policy environment that could make contract farming more effective for smallholders, for example, third-party verification (Saenger et al., 2013).

In the initial title-abstract screening, no studies will be excluded based on the type of study, as the review contains both a meta-analysis and reflection on the impact pathways and enablers and barriers to effectiveness through a case-based comparative analysis. However, the study type and research design will be a main criterion when selecting the quantitative effectiveness studies for meta-analysis (research question 1). The studies selected for meta-analysis also will define the empirical instances for which the enablers and barriers of effectiveness will be compared (research question 2).

Quantitative effectiveness studies

Contract farming is usually initiated by a private company that locates its operations in an area where the product that it wants to buy is widely grown or where agro-ecological conditions are well suited for growing the product of interest. It is common practice for the company to select only those farmers which it believes can offer the right quality at the right price. Large farms in easy-to-reach locations are more likely to be selected. Often, criteria used by firms to select farmers are a combination of observable and non-observable characteristics (Bellemare, 2012). Farmers who are offered a contract can decide whether or not to take it. Those farmers who expect to benefit most from the contract are most likely to accept the offer. Because this decision is at least partly based on unobservable characteristics such as ability, dedication, or knowledge, a self-selection problem is created which might cause bias. Therefore, quantitative effectiveness studies selected for meta-analysis will need to have a design that controls for this selection bias when computing effect estimates. This will include both longitudinal and cross-sectional studies. Following the standard guidelines of systematic review bodies such as the Cochrane and Campbell Collaboration (Higgins and Green, 2011), a hierarchy in methodological rigour in effectiveness studies is applied using the decision tree developed by Waddington et al. (2012).

Some studies use Randomised Control Trials (RCTs) for testing the effectiveness of certain contract components, such as third-party verification (Saenger et al., 2013). However, most studies are expected to have a quasi-experimental design. We include studies using statistical matching (e.g., propensity score matching or PSM, or covariate matching), regression adjustment (e.g., difference-in-differences or DID, and single difference regression analysis, instrumental variables or IV, estimation

and Heckman selection models), as well as other cross-sectional or longitudinal designs which used less rigorous approaches. Given the breadth of designs included, the risk of bias will be assessed using the 3ie's Risk of Bias tool [see chapter 3.5 (Hombrados and Waddington, 2012)].

Qualitative descriptive studies

The quantitative studies alone are expected to give only some of the information needed to identify enablers and barriers for effectiveness in each empirical instance. Therefore, around each example of contract farming (the same instances covered by the studies selected for the meta-analysis), we will search for complementary evidence and background information (see paragraph 3.3). A wide range of documents will be closely examined for potential contribution to this analysis: policy documents, business plans, websites, project initiation documents, etc. This literature will be used not only to reflect in more depth on the relevant barriers and enablers of effectiveness mentioned in each of the effectiveness studies, but also to identify other influencing factors that were not reported yet deemed relevant. Where important information gaps remain, the authors of the selected studies or implementers of the contract farming arrangement may be contacted.

Since the inclusion criteria are potentially large, and this part of the study will include both descriptive information (e.g. programme design and monitoring documents) and analytical literature (e.g. process evaluation documents) we will report transparently the types of evidence obtained for these different types of information. The studies used to make inferences on barriers and enablers will be assessed on relevance and rigor – whether the method used to generate the particular data is credible and trustworthy (Wong et al., 2013) - using the checklist described in paragraph 3.5.

Other Inclusion/ Exclusion Criteria

The search strategy will include electronic data-bases that provide access to non-English literature using English search terms (e.g., www.worldwidescience.org). The research team has a good command of written Spanish, Portuguese, French, English, German and Dutch, which will be used to identify unpublished studies through complementary web-searching. If the language of a study proves to be a problem for data extraction (e.g., Japanese, Chinese), this will be mentioned in the review report.

3.2 Methods Used in Primary Research

To be selected, primary research in quantitative effectiveness studies requires a research design that documents the outcomes of a real functioning contractual arrangement (not the expected/simulated outcomes in newly designed/planned contractual arrangements). Studies of effects need a counterfactual design to critically check the key assumption that these outcomes would not have occurred without the contract farming arrangement. During the process of developing the review, no

study using Randomised Control Trial (RCT) has been identified. Studies that are likely to be included use quasi-experimental designs.

A good example of a study that will be included in the meta-analysis is Miyata et al. (2007), who studied the impact of contracts on growers of apples and green onions in China by comparing contract with non-contract growers. Contract farmers in their sample were selected randomly from lists provided by four private firms; non-contract farmers were selected randomly from lists provided by village leaders. A Heckman model was used to control for sample selection bias, with distance to the village leader as an identifying variable. This identifier was used because village leaders are often involved in selecting farmers for contract schemes.

3.3 Search Methods for Identification of Studies

Search Terms

To find relevant qualitative and quantitative research, we will use a comprehensive search strategy. The retrieved studies will then be screened to identify the core set of studies that give evidence on quantitative effectiveness, used to respond to research question 1, to define the empirical instances of contract farming for meta-analysis. A wider pool of studies will be used for responding to research question 2, to identify relevant information for reflecting on the impact pathways, causal mechanisms, contextual conditions and outcomes of contract farming in these empirical instances. The search strategy detailed in this paragraph concerns the first search process to answer research question 1 and select the empirical instances of contract farming. See Annex 3 for the detailed search strategy used on SCOPUS and CAB-Abstracts.

Electronic Searches

During the preparation of this review protocol, in October 2014, we searched the academic literature through Scopus.com based on provisional research terms. We used the results to fine-tune the search strategy. A more comprehensive electronic search was applied to Scopus and CAB Abstracts in March 2015, with a modified search strategy, following the suggestions of the information retrieval specialist in 3ie, John Eyers. We updated these electronic searches in Scopus.com and CAB Abstracts, in October 2015. In this electronic search we will use all of the following electronic libraries:

- Scopus.com
- CAB Abstracts
- Web of Science
- Agricola
- Econlit
- Tropag & Rural (Royal Tropical Inst, Amsterdam)

Other Searches

In addition to the electronic search, hand-searching and snowballing should add more studies to the review. Additional literature will be searched according to the following four-step process:

Step 1. Non-academic data-bases

First, we will search for complementary academic and non-academic literature in several databases, where electronic search results are impossible or cannot be exported in a useable format. Databases of organisations such as FAO, World Bank and IFAD are searched for grey literature which is not available through the bibliographic databases searched. The list of websites and library locations is still not final but will include at least the following:

- Agris, the FAO database on agricultural science and technology information
- Agecon, research in agricultural and applied research
- Worldscienceweb.org (federation of national science portals where research results are made available by participating nations, providing detailed coverage of global science and research results across language barriers.)
- Eldis (collection of editorially selected and abstracted full-text, online documents on development issues), IDS
- IDEAS (bibliographic database dedicated to Economics and economic research, including Research Papers in Economics database (RePEc)
- Networked Digital Library of Theses and Dissertations
- ProQuest dissertation database
- Best Evidence Encyclopaedia (BEE)
- ESRC (Economic and Social Research Council)
- USAID
- IFPRI
- R4D, DFID
- Jolis: World bank and IMF database, http://external.worldbankimflib.org
- 3ie Database of Impact Assessment, covering impact evaluations conducted in low and middle- income countries, http://www.3ieimpact.org/database_of_impact_ evaluations.html
- Google scholar: general search, the hits will be screened with regard to their potential relevance to the present systematic review.
- Researchgate, https://www.researchgate.net
- SSRN, Social Science Research Network, which includes working papers and submitted papers under review
- FAO Contract Farming Centre. http://www.fao.org/ag/ags/contract-farming/index-cf/en/
- CIAT & CGIAR
- Regoverning Markets, http://www.iied.org/regoverning-markets

Step 2. Snowballing books and review articles

Second, we will use snowballing of references in review articles and books. For example, Oya (2012) provides an extensive overview of the literature on contract farming in Africa focusing on on-going ideological debates and drivers of change. Minot and Ronchi also identified several effectiveness studies (Minot and Ronchi, 2015) and several new volumes have been published that give an overview of research on contract farming (da Silva and Shepherd, 2013; Da Silva and Ranking, 2013). Prowse (2012) analysed 100 studies, mostly grey literature from the FAO database.

Step 3. Citation tracking

Third, citations to the studies used in the meta-analysis will be identified through citation searches (Google Scholar, Scopus, Web of Science) and, if not yet included, screened on the inclusion/ exclusion criteria. Subsequently, the references of all newly included studies will be screened for other relevant studies, using snowballing of references and forward citation tracking.

Step 4. Practitioners' networks

Fourth, the link and information on our dedicated website associated with this systematic review (http://contractfarming-systematicreview.wikispaces.com/) will be circulated among the following communities of practice and e-discussion groups:

- The Knowledge Brokers' Forum (http://www.knowledgebrokersforum.org/);
- Platform on African and European Partnerships in Agricultural Research for Development (http://paepard.blogspot.com/);
- Global Conference on Agricultural and Rural Development-Africa (http://gcardblog.wordpress.com/tag/africa/);
- Global Forum on Rural Advisory Services (http://www.g-fras.org/en/);
- FAO portal on contract farming Contract Farming Resource Centre http://www.fao.org/ag/ags/contract-farming/index-cf/en/
- INNOVAGRO (http://www.redinnovagro.in/)

Step 5. Searching around each empirical instance

Fifth, to answer research question 2, we aim to retrieve additional material that refers to the same contractual arrangement and context which are the focus of the studies selected for meta-analysis. This will allow us to complement the information on context characteristics and contract modalities needed to reflect on the enablers and barriers in the empirical instances of contract farming covered by the studies selected for the meta-analysis. There will be four separate searches around each empirical instance: with the authors name, the name of the contract farming arrangement, the name of contracting firm, and the geographical location of the empirical instance. Next to the data-bases mentioned above in Step 1, we will use Google, www.google.com and LexisNexis Academic http://academic.lexisnexis.nl/. The retrieved documents will be checked for other useful references.

Reference Management and Screening Procedures

The search results used to identify a broad set of potentially relevant qualitative and quantitative empirical studies as well as conceptual papers on contract farming are stored in EPPI Reviewer 4. Each upload will indicate the source and search terms used, by uploading search specific RIS-files).

In the title-abstract screening, reviewers will be over-inclusive, and include all studies that are related to agriculture and in which transactions in product or input markets are mentioned. At least 50% of these references are double-coded. In the full-text screening, all references will be reviewed by at least two reviewers. Two lead reviewers will then independently assess full-text studies for inclusion in the meta-analysis. The lead researcher will reconcile and make a verdict on the final coding.

Data extraction of all studies included in the meta-analysis will be done by two of the lead researchers. Possible disagreements will be arbitrated by a third reviewer.

Title-abstract screening

In the title and abstract screening, studies are excluded if these do not concern or do not comply with one or more of the following six exclusion criteria, mentioned in paragraph 3.1:

1. EXCLUDED no agricultural value chain
2. EXCLUDED forestry (timber, wood)
3. EXCLUDED marine fishery (sea, coastal)
4. EXCLUDED not in low/middle income country
5. EXCLUDED no market related contractual arrangement
6. EXCLUDED not considering farmers
7. EXCLUDED FOR META-ANALYSIS: relevant but no empirical instance
- INCLUDED (might) include empirical instance

Full text screening

After title-abstract screening, we will have a set of publications for which we can retrieve the full text. The references are screened on the above mentioned exclusion criteria about relevance, and using three additional criteria related with our definition of the intervention, contract farming.

Some studies do not assess effectiveness but, for example, differences in characteristics of farmers with and without contracts, or explore differences in performance between contracting farmers. We will use two additional exclusion criteria to further focus on the subset of studies potentially useful for meta-analysis:

8. EXCLUDED no study on impact/ effectiveness
9. EXCLUDED no pre-harvest service delivery

All studies that result from this full-text screening will contain information about one or more empirical instances of contract farming.

Selection for meta-analysis

The studies that remain included will be differentiated based on the methodological rigour of the assessment of their impact.

Data will need to be collected at farm or household level in both intervention and comparison groups. Observational studies are included whenever they control for unobservable characteristics using statistical matching (propensity score or covariate matching) or regression adjustment (difference-in-differences, single difference regression analysis, instrumental variables estimation, and Heckman selection models). Both study designs that collect longitudinal data at baseline and end-line, and those that use cross-sectional end-line data only, are included.

10. EXCLUDED empirical studies without comparison group or other counterfactual design
- INCLUDED empirical studies with comparison group or other counterfactual design

The studies that do not have a counterfactual design will be excluded from meta-analysis. However, a portion of these studies may be retrieved during synthesis if they relate to one or more of the selected empirical instances of contract farming. This non-counterfactual literature may give additional insights in processes or mechanisms related to contract farming, and help to answer our second research question.

3.4 Data Collection and Management

All studies will be entered in the specialist systematic review software – EPPI Reviewer 4 (Thomas et al., 2010). EPPI-Reviewer 4 has been developed and maintained by the EPPI-Centre at the Social Science Research Unit at the Institute of Education, University College London, UK. EPPI-Reviewer is used to screen the studies, to define the empirical instances, to archive the studies used in the qualtitative an qualitative analysis, and for the assessment of study quality and risk of bias.

The meta-analysis will be done by using packages available in Stata. Metan is the main Stata meta-analysis command. Its latest version allows the user to input the cell frequencies from the 2 × 2 table for each study (for binary outcomes), the mean and standard deviation in each group (for numerical outcomes), or the effect estimate and standard error from each study. It provides a comprehensive range of methods for meta-analysis, including inverse-variance–weighted meta-analysis, and creates new variables containing the treatment effect estimate and its standard error for each study. These variables can then be used as inputs into a wide range of other Stata meta-analysis commands (Harris et al., 2010).

The detailed analysis of the qualitative literature around the enablers and barriers in each of the empirical instances will use ATLAS.TI to prepare for qualitative data analysis. ATLAS.TI is more versatile than EPPI Reviewer 4, when coding and synthesising diverse types of evidence to find conditions, drivers, enablers and barriers to effectiveness.

3.5 Critical appraisal and Risk of Bias assessment

The IDCG risk of bias tool (Hombrados and Waddington, 2012) will be used for critical appraisal. As most studies will not have random assignment of the treatment, we will pay particular attention to evaluating the quality of the alternative mechanism for obtaining group equivalence. When propensity score matching (PSM) or Heckman selection equations are used, these should at the very least consider farm size. If relevant for the value chain studied, other farm characteristics determining participation, such as the presence of irrigation equipment, are also required. Distance to roads or distance to town centre are potentially important additional variables. The relevance of any other instruments will be assessed on a case-by-case basis.

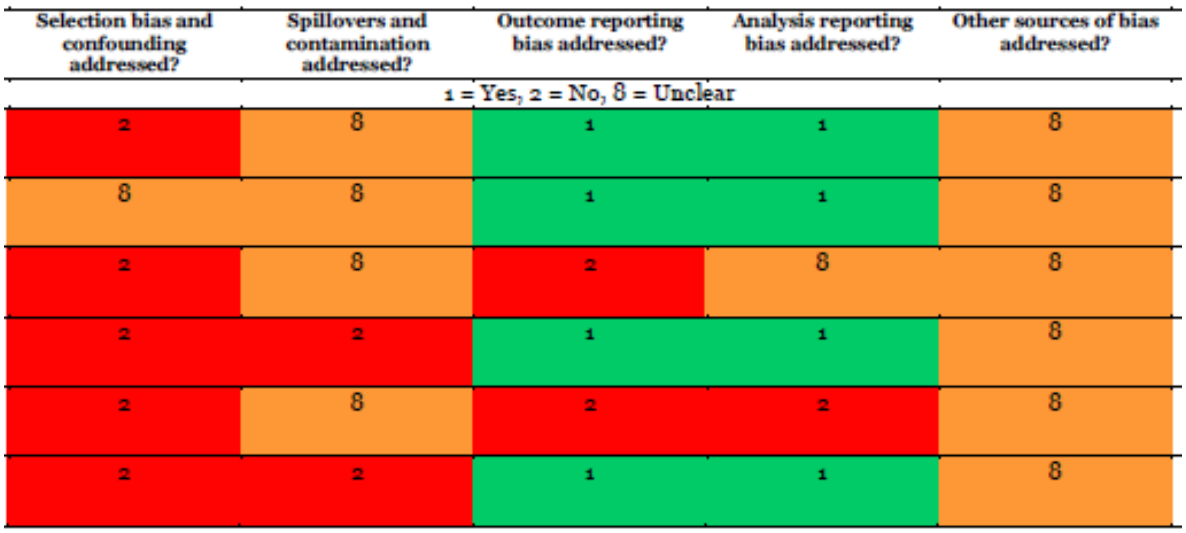
The eight steps in the IDCG risk of bias tool are a critical examination of:

1. Mechanism of assignment
2. Group equivalence
3. Hawthorne and John Henry effects
4. Spill-overs

5. Selective outcome reporting
6. Selective analysis reporting
7. Other bias
8. Confidence intervals

The risk of bias assessment on the quantitative effectiveness studies will be visualised as in Waddington et al. (2014), in Appendix F of the systematic review of Farmer Field Schools, which summarized five (instead of eight) areas where bias may be present.

Figure 5. Risk of bias assessment graph



Source: (Waddington et al., 2014)

Critical Appraisal of Qualitative Studies

Subsequently, we will collect additional information on each empirical instance of contract farming covered in the studies selected for meta-analysis, on enablers and barriers in each empirical instance. We will critically appraise the studies and information sources used to complement the information provided in the quantitative effectiveness studies on the context, enablers and barriers to effectiveness in each of the empirical instances. We will present a table with an assessment of the study quality for all complementary information used in the synthesis to answer the research questions, using the critical appraisal tool used in the Farmer Field School review (Waddington et al., 2014). This tool reviews the presence of sixteen quality attibutes in qualitative studies (see Figure 6).

Figure 6. Critical appraisal tool used in the Farmer Field School review

Research aim	Context	Sampling	Sampling characteristics	Data collection	Data recording	Analysis	Link to relevant lit./ theory	Appropriate methodology	Appropriate sampling	Appropriate methods of data collection	Appropriate analysis	Triangulation	Clarity of analysis and conclusions	Conflict of interest considered/ addressed	Ethical considerations mentioned
1 = Yes, 2 = No, 8 = Partially, 9 = Unclear															
1	1	1	1	1	8	8	1	1	1	1	2	1	8	2	2
1	1	1	1	1	2	2	1	8	8	8	8	1	8	2	2
1	1	1	1	1	2	2	1	8	8	8	2	1	8	2	2
1	1	8	1	1	2	8	1	8	8	8	2	1	2	2	2
1	1	1	1	1	1	1	1	8	8	1	8	1	8	2	1
1	1	1	1	1	8	8	1	8	8	8	8	1	8	2	2
1	1	1	1	1	8	2	8	8	8	8	8	8	1	2	2
1	1	8	1	1	2	8	1	1	8	1	1	1	1	2	1
1	1	1	1	1	1	1	1	1	1	1	8	1	8	2	2

Source: (Waddington et al., 2014)

3.6 Measures of Treatment Effect

The effectiveness studies are expected to measure impact in various ways, with different methods of data-collection and proxy-indicators for different outcome areas. Studies using similar outcome areas will be grouped together in the meta-analysis.

To compare different outcomes measured on a continuous scale, the outcome measure will be converted into standardized mean differences (d), indicating the effect-size in standard deviations between treatment and the control group (Lipsey and Wilson, 2001).

We will use various formulae to calculate SMD. Where information is available, we will use the outcome means in intervention (X_1) and comparison groups (X_2), together with the pooled standard deviation (s_{pooled}), estimated as the weighted average of the standard deviations in intervention group (s_1) and comparison group (s_2).

$$d = \frac{\overline{X}_1 - \overline{X}_2}{s_{pooled}}$$

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}}$$

For studies reporting the result of a t-test between control and treatment group, standardized mean differences will be estimated as:

$$d = t \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

In the unlikely case that only total sample size is reported, it will be assumed that the number of participants is the same in control and treatment group ($n_1 = n_2$).

For studies reporting partial effect sizes - effect sizes reported as regression coefficients from regressions that adjust for other covariates - the beta-coefficient (β) will be used instead of the mean differences (X_1-X_2) in de formula for standardised mean differences (d).

To correct for the fact that some studies have small sample sizes, the standardised effect size for all included studies will be converted into the adjusted mean difference (d') using the following formula:

$$d' = \left[1 - \frac{3}{4N - 9}\right] d$$

Standard deviations of standardized mean differences will be calculated using one of the following formulae:

$$SE = \sqrt{\frac{n_t + n_c}{n_t * n_c} + \frac{SMD^2}{2*(n_t + n_c)}} \quad \text{or} \quad SE = \frac{SMD}{t_\beta}$$

When the outcome variable is not measured on a continuous scale but on a dichotomous scale, the odds-ratio will be used to compare effects. Odds ratios (OR) and corresponding standard errors will be calculated as follows:

$$\text{OR} = \frac{s_t * (n_c - s_c)}{s_c * (n_t - s_t)}$$

$$SE_{OR} = \sqrt{\frac{1}{s_c} + \frac{1}{s_t} + \frac{1}{n_c - s_c} + \frac{1}{n_t - s_t}}$$

where s refers to the number of successful cases in treatment (t) and control group (c) and n to the sizes of the treatment and control group.

Odds ratios may be converted to standardised mean differences using the method suggested by Hasselblad and Hedges (1995):

$$d = Log(OR) \times \frac{\sqrt{3}}{\pi}$$

$$SE_d = \sqrt{V_{Log(OR)} \times \frac{3}{\pi^2}}$$

Conversely, standardised mean differences may be converted to odds ratios using

$$Log(OR) = d\frac{\pi}{\sqrt{3}}$$

$$SE_{Log(OR)} = \sqrt{V_d \frac{\pi^2}{3}}$$

To derive these estimates, information from some studies may be lacking. We will make use of the effect-size calculator (http://www.campbellcollaboration.org/resources/effect_size_input.php). The estimates will be based on the statistical information provided in the studies, or, when missing, on the information provided by the study authors as requested by the review team.

To facilitate interpretation of findings, the estimated pooled standardised effect size could be converted back into the pooled absolute effect size, for example using the approach suggested by Rosenthal and Rubin (1982) who suggest conversion into percentage difference in improvement rate between treatment and control group. This could take the form of a table which converts the pooled standardised effect sizes into values of the outcome variable. In order to do so, assumptions are required regarding baseline values and standard deviations. These assumptions will always be clearly stated, either directly in the table or in the accompanying text.

3.7 Unit of Analysis Issues

The unit of analysis is smallholder farmers in low and middle income countries. We will use the classification used by the World Bank (2014).

In agricultural surveys, the selection of the sample often follows a clustered design. For clustered designs, the assessment of the unit of analysis error will be based on whether the unit of analysis is different from the unit of treatment assignment. If this is the case, the review will assess whether study authors take clustering into account in the analysis (e.g., using multilevel model, variance components analysis or cluster level fixed effects). For those studies with relevant risk of unit of analysis error, corrections will be applied to the standard errors and confidence intervals of those studies. Adjusted standard errors for those studies with relevant risk of unit of analysis error can be estimated as follows (Higgins & Green, 2011):

$$\mathrm{SE}_{\mathrm{corrected}} = \mathrm{SE}_{\mathrm{uncorrected}} * \sqrt{(1 + (\mathrm{m} - 1) * \mathrm{ICC}}$$

Where m is the number of observations per cluster and ICC is the intra-cluster correlation coefficient. The ICC is an estimate of the relative variability within and between clusters. The higher it is, the larger the dependence between the observations within clusters and smaller the effective sample size.

3.8 Criteria for Determining Independent Findings

Most likely, the primary studies used for meta-analysis will provide data on multiple outcome measures, including different proxy-indicators for intermediate and ultimate outcomes. This not a

problem if a study measures outcomes in these different areas and the outcome area can be analysed separately across studies. However, in many cases, multiple proxy-indicators may refer to the same outcome area. In such cases, these outcome measures are taken on the same sample of participants and, therefore, are not independent estimates in the computation of pooled effects because the same study is used multiple times in the same meta-analysis. An explanation of the criteria used to determine whether multiple outcomes from the same or related evaluations are independent data points will be carefully spelled out. We prefer separate stratified analyses, using the proxy-indicator with the lowest risk of bias in the attribution of impact. However, in some cases we may combine groups prior to synthesis, by calculating a 'synthetic' sample-weighted average effect size of the proxy-outcomes that relate to the same outcome area, using appropriate formulae to recalculate variances and standard errors, making covariance assumptions as necessary (Borenstein et al., 2011).

3.9 Dealing with Missing Data

Where data is missing for the meta-analysis of effects of contract farming in the selected empirical instances, authors or implementing agencies will be contacted by e-mail. If key information remains missing, studies will not be included in (all) sub-analyses, but the studies will still be described in the overall analysis.

For the case-based comparative analysis, with the same empirical instances as focal points, the review will extract the relevant available information in the effectiveness studies (see Coding tool), but complement this with other pieces of information about the empirical instance of contract farming retrieved in the complementary targeted search process or received from authors of studies or the implementing agencies. For each empirical instance, we will indicate areas for which information remains incomplete and how this may threaten the validity of the inferences made in the review (Shadish et al., 2002).

3.10 Quantitative Synthesis of Effects

This part of the study is geared to making a map of the available evidence for each step in the result chain, the rigour of the research methods used, and the size of effects that are reported, grouping contractual arrangements that are approximately similar.

Once we have selected the empirical studies which give a quantitative estimate of the size of the effects of contract farming, data will be grouped according to the relevant immediate, intermediate and ultimate outcome area. We will present the standardised effect in each study on each outcome area in forest plots, and in a narrative synthesis we describe the type of outcome and the heterogeneity of the findings. Effectiveness studies of approximately similar contract farming arrangements and similar outcome indicators, likely a small number of studies, will be analysed with pooled effect estimates.

Data Synthesis

Quantitative estimates of effectiveness (effect sizes) in each of the studies will be presented using forest plots and, where possible, synthesised using inverse-variance weighted meta-analysis, using the *metan* package in Stata 13. If meta-analysis is not possible, effect sizes will be presented in forest plots without the pooled effect estimate, and a narrative synthesis will be provided describing the heterogeneity in the findings. We will use the random effects model, acknowledging that the

treatment is not uniform and with context-specific effect-sizes. Random effects meta-analysis assumes that there is some true variation between the included studies, besides random variation due to sampling. The model provides estimates of these two components of variance which can guide further exploration of heterogeneity.

Assessment of Heterogeneity

The presence of heterogeneity will be visually examined by making forest plots of the pooled results and formally tested using the chi-squared test. Given the small number of included studies, the level of significance at which we consider heterogeneity to be present will be set at 0.1, to account for the low power of the test with small samples.

Furthermore, heterogeneity will be assessed with the I-squared statistic, which measures the percentage of variability that is due to real variability between studies. Values close to 100% indicate a large real variability, values close to zero indicate no observable real variability. Final evaluation of test results depend on the significance of the chi-squared test described above and the sign and size of standardised effect sizes.

As we will be using a random effects model for the meta-analysis, between-study heterogeneity can also be assessed with the tau-squared statistic.

Based on our preparatory screening results and discussions with the Advisory Board in February 2015 (Annex 1), we expect that contextual heterogeneity may be an issue that prevents pooled statistics. The most likely causes of this heterogeneity result from the type of product (e.g., perishable and non-perishable products, annual and perennial crops, and intensive animal husbandry) and type of markets (local spot market, processing, local high-value market, export, etc.). The prevalence of heterogeneity depends in part on which products and markets are covered by the identified studies. Also, in the likely absence of a 'standard contract', the heterogeneity of contract clauses and conditions could complicate the drawing of valid inferences regarding the effect size.

To assess the comparability of contractual arrangements between the empirical instances included in the quantitative effectiveness studies, differences in content and conditions will be mapped using the coding tool (see Figure 2 and Annex 3). The extent to which this mapping derives in an adequate control of heterogeneity cannot be answered at protocol stage, and crucially depends on data (study) availability. Based on the number of studies already identified, we expect that the treatment 'contract farming arrangement' might be similar enough to permit generalised inferences based on pooled statistics only for a limited set of studies, probably the studies that relate to annual crops. Where it is not possible to pool information using statistical meta-analysis, we will describe findings narratively and use a theoretical framework of synthesis (e.g. analysis of outcomes along the causal chain). We will also aim to present effect sizes using forest plots.

The case-based comparative analysis will focus on research question 2, and use the studies covered in the meta-analysis. In the case-based comparative analysis we will explore in detail the content and heterogeneity of the contractual arrangements (dispute resolution clauses, payment systems, embedded services, etc.) which are mentioned in the study as enabling conditions that influence the performance of contract farming (enablers, barriers).

Investigation of Heterogeneity

For inclusion in meta-analysis, and especially when pooled statistics are used, the review will ensure that included studies are similar enough that it makes sense to combine them, sensibly divide the included studies into homogeneous groups, or conclude that it did not make sense to combine or group the included studies. Random effects meta-analysis will be employed on each sub-set of studies that are 'approximately comparable', based on our groupings of interventions, exposures and outcomes, drawing on information obtained from the quantitative and qualitative studies. We will code the studies in such a way that these codes can be used as moderator variables in sub-group meta-analysis or meta-regression. We will detail why studies are considered to be sufficiently comparable to permit pooled statistics or not.

Heterogeneity is also relevant in the qualitative synthesis, in which we combine the quantitative effectiveness studies with complementary qualitative information on the same empirical instance of contract farming, to answer research question 2 about enablers and barriers for effectiveness of contract farming. We will assess the research quality that underpinned the findings in these pieces of information. We will follow as much as possible the critical appraisal tool used in the Farmer Field School review (Waddington et al., 2014). The confidence in each aspect will be judged as high, moderate, low, or very low. This assessment will be described and justified in a summary table of all pieces of qualitative information used.

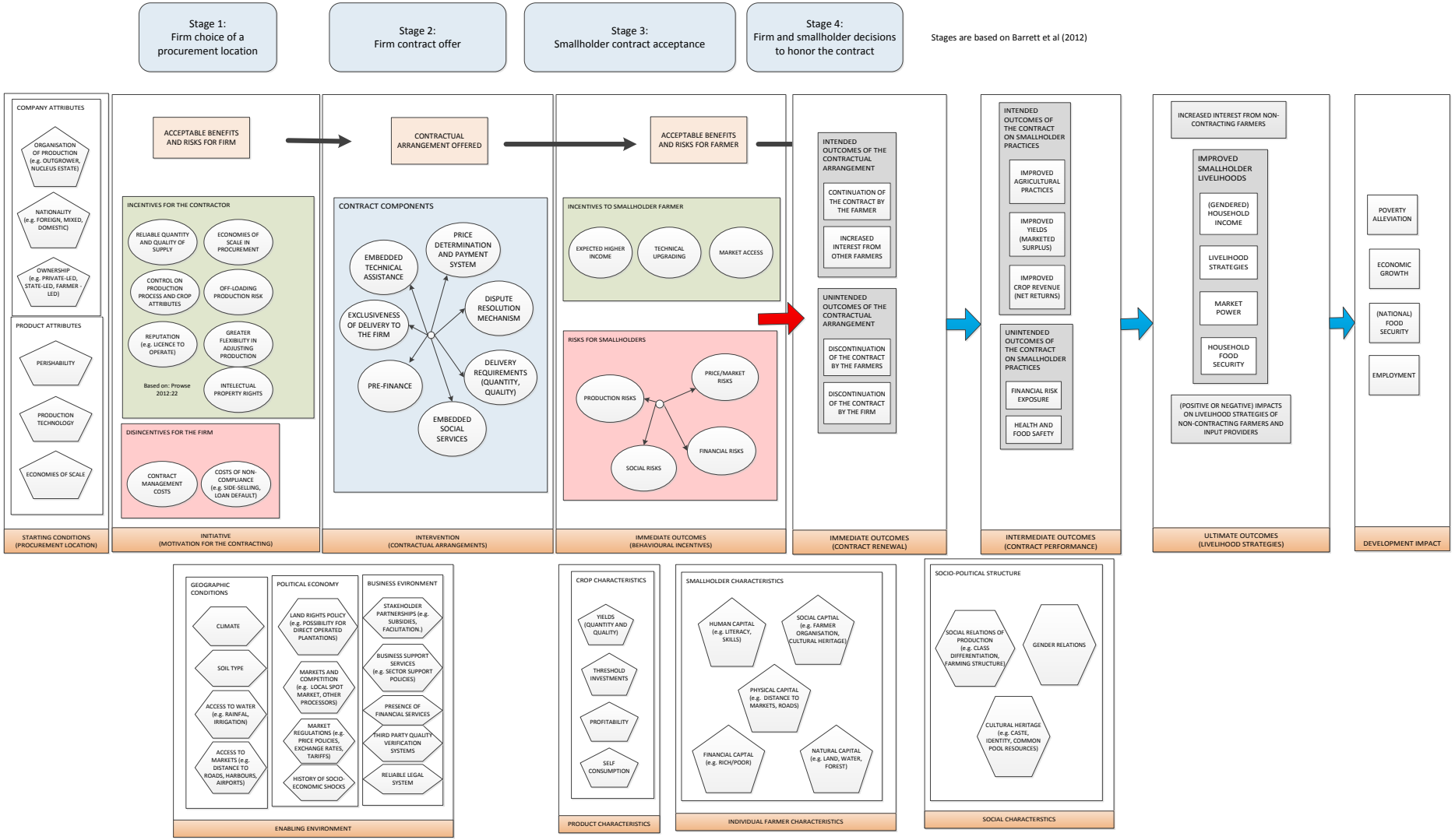
Sensitivity Analysis

Whenever we compute effect sizes with pooled statistics, a sensitivity analysis will be performed to check the robustness of these findings. Therefore, any overall pooled results will be compared to pooled-results of sub-sets of studies. For example, one of the sub-set analyses will be based on the study design and methodology used in the included studies. Studies will also be grouped according to their overall risk of bias scores and compared. We also expect to compare studies based on the period over which they measure the impact, i.e., group studies based on the number of years farmers have been receiving the contract farming arrangement at the moment of measurement. During the review process, other relevant comparisons are likely to be found, and will be included in the sensitivity analyses.

Assessment of Reporting Bias

We will analyse the quantitative effectiveness studies using funnel plots to visually explore publication bias. In addition, we will report Egger's meta-regression test (Egger et al., 1997) when there are sufficient studies to be able to do so, using the Stata *metabias* command.

Figure 7. Conceptual framework use to code the information to answer research question 2 on barriers and enablers.



3.11 Case-based Comparative Analysis of Enabling Conditions

This second research question concerns conditions that work as enablers or barriers for the positive effects to materialize. It consists of two sub-questions: 2.1 What are the enabling and disabling conditions that explain the effectiveness and sustainability of the contract farming arrangements for smallholder famers?, and 2.2 Are there configurations of conditions that may predict success or failure of contract farming arrangements? We anticipate that the quantitative studies alone are unlikely to provide all information needed to assess the presence or absence and relevance or irrelevance of these conditions. Therefore, we will complement the impact studies with qualitative studies and other pieces of information that may shed light on the conditions that are relevant in the context of each empirical instance. The studies selected for the quantitative studies (meta-analysis) will define the specific *instances* of contract farming that will be covered in this qualitative analysis, i.e., the specific places and situations described in these studies. For each of these instances, complementary literature will be identified using a combination of snowballing and targeted searches. Internet searches will be used to identify pieces of information that can inform us about the particular conditions and dynamics in each empirical instance, especially when not reported with sufficient detail in the effectiveness studies (see step 5 in Chapter 3.3).

In Figure 7, we identify three sets of enablers and barriers.

- geographical conditions (climate, soil type, access to water, access to markets),
- political/ economic issues (land rights, market competition, market regulations, history of economic shocks), and
- characteristics of the business environment (stakeholder partnerships, business support services, financial services, third party quality verification systems, reliable legal system).

We will check the relevance of these enabling conditions for each empirical instance of contract farming. This mapping will be used to draw generalised inferences about enablers and barriers to the successful implementation of contract farming with smallholders. In doing so, we use the evidence to build more refined impact pathways than the one depicted in Figure 2,. The resulting typology and related refined impact pathways will be discussed with members of the advisory board, to check its face validity, before being presented in the review report.

If deemed necessary - in consultation with the Advisory Board and 3ie - this part of the systematic review may have a more limited focus and concentrate on the subset of contractual arrangements for which findings about enabling conditions have fewer threats to external validity. For example, focusing especially on the empirical instances of contract farming identified in the meta-analysis that relate to a key moderating variable that is common to a subset of studies, e.g., related to annual agricultural crops (see Annex 1 – Minutes of the Advisory Board meeting – February 2015).

3.12 Validity Threats to the Results

We anticipate that we will not be able to draw strong conclusions on the effectiveness of contract farming because of the expected wide heterogeneity of contractual arrangements, contexts, limited number of replications of studies that refer to the same empirical instance, and the quality of the study designs. The results of the review will be more explorative in nature and will suggest areas to be

explored in subsequent empirical research on the effectiveness of contract farming for smallholder income and food security.

4. Plans for Updating the Review

Funding permitting, the review will be updated within three years, based on an electronic search for new academic studies, complemented with a focused search in the institutes and data-bases mentioned in review. The newly retrieved quantitative effectiveness studies will be assessed using the same coding tool as in this review.

5. Time frame

Searches for published and unpublished studies	09/2014—11/2015
Screening/ relevance assessments	10/2014—12/2015
Pilot testing of study codes and data collection	10/2015—11/2015
Extraction of data from research reports	12/2015—02/2016
Statistical synthesis	11/2015—12/2015
Qualitative synthesis (if appropriate)	12/2015—03/2016
Preparation of report	03/2015—04/2015
Report revisions	05/2015—10/2016
Presentation of findings/ policy influence work	10/2016—12/2016

6. Acknowledgments

Research discussed in this publication has been funded by the International Initiative for Impact Evaluation, Inc. (3ie), provided under Systematic Review – 6 Window. Additional support was provided by LEI Wageningen UR and Ghent University.

7. The Review Team and Contributions of Authors

We have opted for a small core team, consisting of Giel Ton from LEI Wageningen UR and Wytse Vellema, Sam Desiere and Marijke D'Haese from Ghent University, to make internal coordination as effective as possible. The proximity of Wageningen and Ghent allows team members and students to have regular face-to-face meetings, making it easier to maintain internal consistency. All core team members are specialised in research on smallholder farmers in developing countries. Searching will be done with the help of a senior librarian at LEI Wageningen UR. Easily definable tasks, such as title-abstract screening and full-text mapping, will be delegated to several graduate students and interns, including Sophia Weituschat, although they will always be double coded in EPPI Reviewer by one of the lead researchers.

Giel Ton

Giel Ton works as Senior Researcher at LEI Wageningen UR, The Netherlands, and has ample experience in managing project teams, especially to align methodologies and approaches between different research paradigms (Ton, 2012). He has been involved in the design and monitoring of several contractual arrangements in agriculture, especially related to product certification, such as Fair Trade, Utz Certified and Rainforest Alliance (Waarts et al., 2012), and research on internal organizational management of contract farming (Ton and Mheen-Sluijer, 2009) and collective

marketing (Ton et al., 2014a). In this context, he has experience in developing impact logic frameworks to evaluate interventions. He has been involved in two systematic literature reviews that have been published (Ton et al., 2013). He also participated in the online 'Elluminate course for review teams' by EPPI-Centre in 2011 and the Induction Course by 3ie in November 2014. One review concerned the impact of innovation grants that are directed to smallholder farmers (Ton et al., 2013; Ton et al., 2015). A second systematic review was commissioned by the Dutch Evaluation Office (IOB) and explored evidence on the effectiveness of policies to enhance Corporate Social Responsibility (CSR). Giel Ton participated in the design of the research (protocol), especially in the refinement of conceptual framework and intervention logic to map the literature in a more homogenous typology (Ingram et al., 2012). Giel Ton will lead the development of the review methodologies and do the qualitative synthesis. He has experience with Atlas.ti for qualitative data synthesis from his PhD-study on the impact of small-grant funds on economic farmer organisation.

Wytse Vellema

Wytse obtained his PhD at the department of agricultural economics at Ghent University, Belgium. His research focuses on how different type of sourcing arrangements affect the inclusion and well-being of smallholder farmers. He is lead researcher on the impact evaluation team for a development intervention by CRS. The project aims to improve the livelihoods, income, and food security of smallholder coffee farmers in the Southern Colombian Andes and the Ecuadorian Amazon through improved market access and increased on-farm diversification. In the context of this assignment, he has gained experience with developing impact logic frameworks and measurement tools for monitoring and evaluation. He has a background in international economics and business, specialising in econometrics. During one year at Cambridge University he further developed his theoretical and quantitative skills. After a short detour in mergers and acquisitions as a valuations specialist, he finished a research master's degree in rural development. During this degree study, he conducted field work for a randomized controlled field experiment, on which he later wrote his thesis. In his PhD work, he used new institutional economics and transaction cost theory to research smallholder sourcing arrangements, focusing in particular on the factors driving the specific form of model chosen, the success of these models, and how they reach different groups of farmers and affect their wellbeing.

Marijke D'Haese

Marijke has lived or worked in Africa most her life, and specializes in the interaction between poverty, food security, and the role of market institutions. Most of her work analyses decision-making by rural households as farmers of food and cash crops as well as producers or collectors of non-timber forest products or traders, testing micro-economic models with primary data. Although she does not have previous experience with meta-analysis, she has extensive experience with applied econometrics and statistics, and is fluent in Stata.

Sam Desiere

Sam obtained his PhD at the department of agricultural economics at Ghent University. He holds a master's degree in economics from the same university. His research interests include food security and rural development in developing countries, particularly in sub-Saharan Africa. Recently, he focused on developing and evaluating indicators to quantify poverty and food security (Desiere et al., 2015a; Desiere et al., 2015b). This experience is valuable for the framework of this systematic review,

as assessing the reliability of outcome variables is an essential step when conducting a careful meta-analysis. In addition to his research on indicators, he is involved in a research project about the agricultural sector in Burundi. In this context, he has gained experience with critically analysing data from nationally representative, agricultural surveys. In the framework of this assignment, Sam Desiere will contribute to the meta-analysis part of the review.

Sophia Weituschat

Sophia Weituschat was born and raised in Berlin, Germany. In 2011, after completing a year at Humboldt University, she started to study International Development at Wageningen University. She is currently specializing in an MSc in development economics and works as a student assistant.

Ans Brouwer

Ans is information specialist Plant Sciences at Wageningen UR Library. She studied Plant sciences at Wageningen University. She has more than 30 years of experience in information and collection management, teaching information literacy (including EndNote) and performing literature search including extensive literature searches for systematic reviews. Her research fields include information retrieval, plant sciences in general, plant pathology, plant protection and plant breeding in particular..

8. Statement Concerning Conflict of Interest

All authors have published on the topic of smallholder market access. Most of their studies were qualitative in nature and have been used for building the conceptual framework. Marijke D'Haese wrote some studies that are likely to be included for meta-analysis. The eligibility and rigour of these studies will be assessed by the other reviewers and checked with the external advisers, part of the Advisory Board of the review. There are no financial or similar conflicts of interest.

9. References

Barrett CB, Bachke ME, Bellemare MF, et al. (2012) Smallholder participation in contract farming: Comparative evidence from five countries. *World Development* 40: 715-730.

Bellemare MF. (2012) As You Sow, So Shall You Reap: The Welfare Impacts of Contract Farming. *World Development* 40: 1418-1434.

Bellemare MF and Novak L. (2015) Contract Farming and Food Security.

Borenstein M, Hedges LV, Higgins JP, et al. (2011) *Introduction to meta-analysis*: John Wiley & Sons.

da Silva C and Shepherd A. (2013) Editorial: The multi-faceted dimensions of contract farming. *Food Chain* 3: 127-130.

Da Silva CA and Ranking M. (2013) Contract farming for inclusive market access. Rome: FAO.

Desiere S, D'Haese M and Niragira S. (2015a) Assessing the cross-sectional and inter-temporal validity of the Household Food Insecurity Access Scale (HFIAS) in Burundi. *Public Health Nutrition* FirstView: 1-11.

Desiere S, Vellema W and D'Haese M. (2015b) A validity assessment of the Progress out of Poverty Index (PPI)™. *Evaluation and program planning* 49: 10-18.

Eaton C and Shepherd A. (2001) *Contract farming: partnerships for growth*: Food & Agriculture Org.

FAO IFAD WFP. (2014) The state of food insecurity in the world. Rome.

Harris R, Bradburn M, Deeks J, et al. (2010) METAN: Stata module for fixed and random effects meta-analysis. *Statistical Software Components*.

Hasselblad V and Hedges LV. (1995) Meta-analysis of screening and diagnostic tests. *Psychological Bulletin*: 167-178.

Heady D and Ecker O. (2013) Rethinking the measurement of food security: From first principles to best practice. *Food Security* 5: 327-343.

Higgins J and Green S. (2011) Cochrane handbook for systematic reviews of interventions Version 5.1. 0 *www. cochrane-handbook. org.* The Cochrane Collaboration.

Hombrados J and Waddington H. (2012) IDCG Risk of Bias Tool. London.

Ingram V, Ton G, de Grip K, et al. (2012) Protocol for the Systematic Review on the effects of Corporate Social Responsibility.

Jia X and Bijman J. (2013) Contract farming: Synthetic themes for linking farmers to demanding markets. In: Da Silva C and Ranking M (eds) *Contract Farming for Inclusive Market Access.* Rome: FAO, 21-38.

Lindholm A. (2014) The (re) emergence of contract farming in Sub-Saharan Africa: moving from land grab to power grab?

Lipsey MW and Wilson DB. (2001) *Practical meta-analysis*: Sage publications Thousand Oaks, CA.

Little PD and Watts M. (1994) *Living under contract: contract farming and agrarian transformation in sub-Saharan Africa*: Univ of Wisconsin Press.

Minot N. (1986) Contract farming and its effect on small farmers in less developed countries. Michigan State University, Department of Agricultural, Food, and Resource Economics.

Minot N and Ronchi L. (2015) Contract farming: Risks and benefits of partnership between farmers and firms.

Minten B, Randrianarison L and Swinnen JF. (2009) Global retail chains and poor farmers: Evidence from Madagascar. *World Development* 37: 1728-1741.

Miyata S, Minot N and Hu D. (2007) Impact of contract farming on income. Rep. International Food Policy Research Institute. Print.

Narayanan S. (2013) Smallholder attrition in contract farming schemes in India: extent, causes, and concerns. *Food Chain* 3: 155-170.

Oya C. (2012) Contract Farming in Sub-Saharan Africa: A Survey of Approaches, Debates and Issues. *Journal of Agrarian Change* 12: 1-33.

Oya C. (2015) Effects of Certification Systems for Agricultural Commodity Production on Socio-economic Outcomes of Beneficiaries in Low-and Middle-Income Countries: a Systematic Review.

Prowse M. (2012) Contract Farming in Developing Countries - a review. *A Savoir.* Antwerpen.

Rosenthal R and Rubin DB. (1982) A simple, general purpose display of magnitude of experimental effect. *Journal of educational psychology* 74: 166.

Saenger C, Torero M and Qaim M. (2013) Impact of Third-Party Enforcement of Contracts in Agricultural Markets–A Field Experiment in Vietnam. GlobalFood Discussion Papers.

Setboonsarng S and Leung P. (2014) Making Globalization Work Better for the Poor through Contract Farming.

Shadish WR, Cook TD and Campbell DT. (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*: Houghton Mifflin Co. Boston, MA.

Sivramkrishna S and Jyotishi A. (2008) Monopsonistic exploitation in contract farming: articulating a strategy for grower cooperation. *Journal of International Development* 20: 280-296.

Stewart R, Erasmus Y, Zaranyika H, et al. (2014) The Effects of Training, Innovation and New Technology on African Smallholder Farmers' Wealth and Food Security: A Systematic Review.

Thomas J, Brunton J and Graziosi S. (2010) EPPI-Reviewer 4.0: software for research synthesis. *EPPI-Centre Software. London: Social Science Research Unit, Institute of Education*.

Ton G. (2012) The Mixing of Methods: a three-step process for improving rigour in impact evaluations. *Evaluation* 18: 5-25.

Ton G, Flores L, Monasterios R, et al. (2014a) Capabilities and Performance in Collective Marketing: the importance of learning to cope with agency dilemmas. In: Christy R, da Silva C, Mhlanga N, et al. (eds) *Innovative Institutions, Public Policies and Private Strategies for Agro-Enterprise Development.* FAO / World Scientific Publishing Co. Inc., 113-149.

Ton G, Grip Kd, Klerkx L, et al. (2013) Effectiveness of Innovation grants to smallholder agricultural producers: an explorative systematic review. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Ton G, Klerkx L, de Grip K, et al. (2015) Innovation grants to smallholder farmers: Revisiting the key assumptions in the impact pathways. *Food Policy* 51: 9-23.

Ton G and Mheen-Sluijer J. (2009) Contract Farming Checklist: a tool for reflection on critical issues in contract farming arrangements in developing countries. Wageningen: LEI Wageningen - ICCO - EVD.

Ton G, Vellema S and Ge L. (2014b) The Triviality of Measuring Ultimate Outcomes: Acknowledging the Span of Direct Influence. *IDS Bulletin* 45: 37-48.

United Nations Population Division. (2014) *World Urbanization Prospects: The 2014 Revision*: United Nations Publications.

Vellema S. (2002) *Making Contract Farming Work? Society and technology in Philippine transnational agribusiness,* Maastricht: Shaker Publishing.

von Hagen O and Alvarez G. (2011) The Impacts of Private Standards on Global Value Chains. *Literature Review Series on the Impacts of Private Standards, Part I.* Geneva: International Trade Centre.

Vorley B and Proctor F. (2008) Inclusive business in agrifood markets: evidence and action.

Waarts Y, Ge L, Ton G, et al. (2012) Sustainable tea production in Kenya: Impact assessment of Rainforest Alliance and Farmer Field School training. *LEI Wageningen UR, The Hague*.

Waddington H, Snilstveit B, Hombrados JG, et al. (2014) Farmer Field Schools for Improving Farming Practices and Farmer Outcomes in Low- and Middle-income Countries: A Systematic Review. *Campbell Systematic Reviews.*

Waddington H, White H, Snilstveit B, et al. (2012) How to do a good systematic review of effects in international development: a tool kit. *Journal of Development Effectiveness* 4: 359-387.

World Bank. (2007) *World Development Report 2008: Agriculture for Development,* Washington: World Bank.

Annex 1 – Minutes Advisory Board

Advisory board meeting "protocol"

13 February 2015

Introduction

Attending:

- Dave Boselie of the Sustainable Trade Platform (IDH), a public-private partnership which co-invests in sustainable sourcing.
- John Hellin of CYMMIT, who does agricultural development research on contract farming with smallholders, particularly in, but certainly not limited to, cereals.
- Andrew Shepherd, currently consultant, ex-FAO and CTA, "the Shepherd in Eaton and Shepherd (2001)", with broad policy and writing experience on contract farming.
- Jane Njuguna, AGRA, one of the donors of the study, working with smallholder farmers in 18 African countries, taking a value chain approach to poverty alleviation.
- Ruerd Ruben, Research Coordinator Food Security, Value Chains & Impact Analysis at LEI - Wageningen University, impact of value chains on food security.
- Martina Vojtkova, who is responsible for this systematic review in 3ie and will provide technical and methodological support.
- Nicoliene Oudwater from DFID, one of the donors of the study, who is interested in how value chain interventions such as contract farming affect uptake of technology
- Wytse Vellema, project researcher, is a PhD candidate at Ghent University studying smallholder integration into global value chains, focusing on business models and impact.
- Giel Ton, project leader, based at LEI, is experienced in impact evaluations and farmer organizations specializing in Latin America.

Excused:

- Marijke D'Haese, project researcher, professor Rural Development Economics at Ghent University.

In brief

After a brief round of introductions (see above), the meeting was held in two rounds. In the first round, remaining issues after adoption of the comments of each board member in the latest version of the protocol were discussed. This discussion focused on the definition and measurement of smallholders. Some refinements were suggested and adopted. In the second round, the scope of the qualitative synthesis was discussed. Given the heterogeneity among contracts, products, farmers, and contexts, the decision was made to focus on annual crops, distinguishing between perishable and non-perishable products. What follows are the more detailed minutes of the meeting.

Minutes of the meeting

Protocol

Dave Boselie stressed the importance of the profile of farmers. It is important to distinguish between different typologies or clusters of farmers, and their characteristics compared to other farmers, as this is expected to affect both the type of contract as well as its impact.
John Hellin made two points. First, that the livelihood diversity should be taken into account, particularly the question to what extent contract farming generates opportunities in uptake of non-

agricultural activities and investments. Second, potential risk mitigation of contracts, mentioning the example of including crop index insurance in contracts.
Andrew Shepherd discussed the scope of the review. As most farmers in Africa are smallholder farmers, he rejected the concept of using a maximum number of hectares to define a farmer as a smallholder. Only large commercial estates should be excluded. Furthermore, it is important to bear in mind that most businesses do not focus on the poorest of the poor.
Jane Njuguna concurred with Andrew that land size is not a valid criterion, as this differs markedly between countries and contexts. In her experience, not only the contract itself but also the way farmers are organized matters: group versus individual contracts. As farmers often self-select, it is often the more commercially oriented farmers who seek market access through contracts. Finally, the size of the labour commitment is often substantial. When information is available, this should be included in the evaluation.
Ruerd Ruben was happy to see the change from household income to net household income in the protocol, as substitution effects between activities are likely to be substantial. An important additional effect is the certainty or reduced variability of income. Furthermore, he noted that different products have different contracts, giving the example of perishable products such as vegetables and milk. A final consideration, understandably outside of the scope, is the firm side of the contract.
Martina Vojtkova suggested we think critically about how to operationalize the definition of smallholders and potential subgroups or clusters. The same should be done for outcomes. A primary outcome is an outcome a study is required to contain in order to be included in the review. Other outcomes are complementary, and can be used to provide additional information. The heterogeneity between contracts, products, farmers, and contexts can be accounted for with random effects meta-analysis.
Nicoliene Oudwater concurred with John on the importance of mapping the impact of contracts on livelihood opportunities outside of farming. The "opportunity to diversify" is an important outcome in itself, even without direct income effects. Important not to exclude non-agricultural activities or investment in education.

Follow-up actions
The discussion focused on two points: farmers and outcomes. Further refinement of the definition of smallholder farmers is important, taking into account the type of farmers and recognising heterogeneity between groups of farmers. However, this definition will not be used ex-ante to exclude studies from the review. Rather, the definitions used in each paper will be used. The review will look how this is done in each of the studies selected for meta-analysis and will use the four criteria used by Steward (mentioned in the protocol) to map how this is done in the studies. Furthermore, the endogeneity between contract, product, and type of farmers should be taken into account. The changes in time in characteristics of contracted farmers will be discussed in detail. Organization of farmer support (groups, cooperatives, associations), present in many contract farming arrangements, should be taken into account.
Several suggestions were made to specify relevant outcomes. Rather than focus on net income from the contracted product, the focus should be on net income effects, as this accounts for possible substitution effects. As contracts have an important risk mitigation function, sometimes including specific insurance clauses, their effect on the certainty of income should be considered. Another important consideration is the labour commitment required. Finally, the importance of contracts for generating income-generating opportunities outside of agriculture are of particular interest, as it allows mapping livelihood trajectories.
Finally, heterogeneity between farmers, contracts, contract organization, products was mentioned at various points in the discussion. For the meta-analysis, this heterogeneity will be illustrated in the meta-analysis, taking into account the sources of heterogeneity mentioned in the discussion, i.e. farmer type, crop insurance, group or individual farmer organisation, and contract and product

characteristics. Only when the treatment and outcome variables are sufficiently comparable, pooled random effects will be presented in the meta analysis output.

Qualitative synthesis

Topic was introduced by Giel, who posited the question of how to get most relevant information out of the qualitative part of the synthesis, i.e. the optimal balance between breadth and depth. Given the heterogeneity between contracts and the endogeneity of contract type with crop type, a possible solution would be to focus on only one product group. Fruits and vegetables, the largest group, would then be the most likely option.
Dave Boselie initially dropped out of the conversation but was able to join again while Andrew was speaking. After listening to the others, he understood the rationale and agreed with the suggestions made.
John Hellin agreed that to obtain sufficient depth and maintain comparability, focusing on a more uniform type of contractual arrangements would make sense. Fruits and vegetable contracts were the obvious candidate.
Andrew Shepherd suggested to focus on annual crops, rather than fruits and vegetables. Annual crops are a relevant group, because they are easy to move into and out of, which makes it better possible to assess if the contractual arrangements serves smallholders or not. Perennials require a completely different relationship between company and farmer.
Jane Njuguna missed her turn due to connection problems and later concurred with the suggestions made.
Ruerd Ruben concurred with Andrew on the difference between annual and perennial crops, and suggested that out of the two, annual crops were by far the more interesting. Within the annual crops, a distinction should be made between perishable and non-perishable products, as this characteristic strongly influences contract conditions.
Martina Vojtkova made three main points. First, that a good way to focus the qualitative synthesis would be to focus on the same instances to which also the quantitative studies referred to. In other words, to only use the qualitative literature referring directly to projects or contracts which are part of the meta-analysis. This makes the qualitative studies support the main research question on effectiveness. Second, that a general framework could be build using the information from the broader set of qualitative and quantitative studies. This framework could provide a comparison for the information obtained on each empirical instances. Third, selection of qualitative studies to include for building the framework should be based on some predetermined criteria, to make sure low quality studies were excluded from contributing to the framework.
Nicoliene Oudwater had no further additions to what had already been said, and agreed with the general approach discussed.

Follow-up actions
To gain sufficient depth and ensure comparability between studies, the qualitative synthesis will focus on contracts for annual crops. A distinction will be made between perishable and non-perishable products. To answer the second research question, the team will search for complementary qualitative and quantitative studies that directly relate to one of the empirical instances included in the meta-analysis. When used to infer about barriers and enablers for effectiveness, the quality of each of these studies will be assessed. More general studies and conceptual literature will be used but without necessarily resulting from a systematic screening of all possible qualitative studies. The process will thus be different to the systematic search and screening process to derive the quantitative studies for the meta analysis).

General comments

A final round of general comments was held to give everybody an opportunity to discuss topics not yet or insufficiently emphasized during the rest of the meeting.
Dave Boselie discussed the cases that IDH was currently involved in. Two of these cases, in Ethiopia and South Africa, were considered to be of particular interest. It was agreed to share relevant information.
John Hellin expressed his interest in research resulting from the project, and asked about what more was expected from the advisory board. A brief timeline is given in the conclusion below.
Andrew Shepherd wondered to what extent third-party arrangements like cooperatives or third-party input provision were covered by the review. The answer was that the primary criterion to include or exclude a study would be the existence of a direct relation between the contract with the firm and the services made available to the farmer. When, next to material or financial resources third-party services like certification are present, these will be considered as contract farming.
Jane Njuguna had no further comments, indicating her immediate concerns were met.
Ruerd Ruben stressed the relevance of the history of the contracts in each empirical instance. Contracts are not standard, but change over time. The contract type is thus endogenous. He also stresses that the systematic review is not 'classical', because the treatment is extremely heterogeneous and hardly exogenous.
Martina Vojtkova discussed the timeline of protocol approval and suggested to use the guidelines she emailed Giel as much as possible, as these contained the criteria used by the review committee. John, the search specialist, has been contacted and promised to provide feedback on the search strategy.
Nicoliene Oudwater had to leave the call earlier.

Follow-up actions
The revised version of the protocol is expected to be submitted to 3ie in the week of 16 February 2015. The review committee in 3ie is already lined up and feedback will be given as soon as possible, including detailed feedback on the search strategy. Hopefully the protocol will be accepted in March. If additional searches are necessary, they will be done as soon as possible after protocol acceptance. In the meantime, authors of already identified studies will be contacted to obtain data required for the meta-analysis. The team will also contact Dave for information on on-going contract farming projects. In May the meta-analysis is expected to start. An email will be sent to the advisory board members to inform about the selected empirical instances; no detailed feedback is required at this stage. In September a first draft synthesis report is expected to be ready. Around that time a feedback session and meeting will be planned with the advisory board.

Annex 2 – Search strategy

Elaborated with the kind assistance of John Eyers (3ie)

Scopus search – Searched 30th September 2015

((((TITLE-ABS-KEY ((afghanistan OR albania OR algeria OR angola OR argentina OR armenia OR armenian OR aruba OR azerbaijan OR bangladesh OR benin OR byelarus OR byelorussian OR belarus OR belorussian OR belorussia OR belize OR bhutan OR bolivia OR bosnia OR herzegovina OR hercegovina OR botswana OR brasil OR brazil OR bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR "Upper Volta" OR burundi OR urundi OR cambodia OR "Khmer Republic" OR kampuchea OR cameroon OR cameroons OR cameron OR camerons OR "Cape Verde" OR "Central African Republic" OR chad OR china OR colombia OR comoros OR "Comoro Islands" OR comores OR mayotte OR congo OR zaire OR "Costa Rica*" OR "Cote d'Ivoire" OR "Ivory Coast" OR cuba OR djibouti OR "French Somaliland" OR dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR ecuador OR egypt OR "United Arab Republic" OR "El Salvador" OR eritrea OR ethiopia OR fiji OR gabon OR "Gabonese Republic" OR gambia OR gaza OR "Georgia Republic" OR "Georgian Republic" OR ghana OR grenada OR guatemala OR guinea OR guiana OR guyana OR haiti OR hungary OR honduras OR india OR maldives OR indonesia OR iran OR iraq OR jamaica OR jordan OR kazakhstan OR kazakh OR kenya OR kiribati OR korea OR kosovo OR kyrgyzstan OR kirghizia OR "Kyrgyz Republic" OR kirghiz OR kirgizstan OR "Lao PDR" OR laos OR lebanon OR lesotho OR basutoland OR liberia OR libya OR macedonia OR madagascar OR "Malagasy Republic" OR malaysia OR malaya OR malay OR sabah OR sarawak OR malawi OR mali OR "Marshall Islands" OR mauritania OR mauritius OR "Agalega Islands" OR mexico OR micronesia OR "Middle East" OR moldova OR moldovia OR moldovian OR mongolia OR montenegro OR morocco OR ifni OR mozambique OR myanmar OR myanma OR burma OR namibia OR nepal OR "Netherlands Antilles" OR "New Caledonia" OR nicaragua OR niger OR nigeria OR pakistan OR palau OR palestine OR panama OR paraguay OR peru OR philippines OR philipines OR phillipines OR phillippines OR "Puerto Ric*" OR romania OR rumania OR roumania OR rwanda OR ruanda OR "Saint Lucia" OR "St Lucia" OR "Saint Vincent" OR "St Vincent" OR grenadines OR samoa OR "Samoan Islands" OR "Navigator Island" OR "Navigator Islands" OR "Sao Tome" OR senegal OR serbia OR montenegro OR seychelles OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR somalia OR "South Africa" OR sudan OR suriname OR surinam OR swaziland OR syria OR tajikistan OR tadzhikistan OR tadjikistan OR tadzhik OR tanzania OR thailand OR togo OR togolese republic OR tonga OR tunisia OR turkey OR turkmenistan OR turkmen OR uganda OR ukraine OR uzbekistan OR uzbek OR vanuatu OR "New Hebrides"

OR venezuela OR vietnam OR "Viet Nam" OR "West Bank" OR yemen OR yugoslavia OR zambia OR zimbabwe))) OR (TITLE-ABS-KEY ("Developing Countries" OR africa OR asia OR caribbean OR "West Indies" OR "South America" OR "Latin America" OR "Central America" OR ((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income" OR underserved OR "under served" OR deprived OR poor*) W/1 (countr* OR nation* OR population* OR world)))) OR (TITLE-ABS-KEY (((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income") W/1 (economy OR economies)) OR (low* W/1 (gdp OR gnp OR "gross domestic" OR "gross national")) OR (low W/3 middle W/3 countr*))) OR (TITLE-ABS-KEY (((lmic OR lmics OR "third world" OR "lami countr*")) OR "transitional countr*"))) AND (TITLE-ABS-KEY ((contract* W/2 farm*))))) OR ((((TITLE-ABS-KEY ((afghanistan OR albania OR algeria OR angola OR argentina OR armenia OR armenian OR aruba OR azerbaijan OR bangladesh OR benin OR byelarus OR byelorussian OR belarus OR belorussian OR belorussia OR belize OR bhutan OR bolivia OR bosnia OR herzegovina OR hercegovina OR botswana OR brasil OR brazil OR bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR "Upper Volta" OR burundi OR urundi OR cambodia OR "Khmer Republic" OR kampuchea OR cameroon OR cameroons OR cameron OR camerons OR "Cape Verde" OR "Central African Republic" OR chad OR china OR colombia OR comoros OR "Comoro Islands" OR comores OR mayotte OR congo OR zaire OR "Costa Rica*" OR "Cote d'Ivoire" OR "Ivory Coast" OR cuba OR djibouti OR "French Somaliland" OR dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR ecuador OR egypt OR "United Arab Republic" OR "El Salvador" OR eritrea OR ethiopia OR fiji OR gabon OR "Gabonese Republic" OR gambia OR gaza OR "Georgia Republic" OR "Georgian Republic" OR ghana OR grenada OR guatemala OR guinea OR guiana OR guyana OR haiti OR hungary OR honduras OR india OR maldives OR indonesia OR iran OR iraq OR jamaica OR jordan OR kazakhstan OR kazakh OR kenya OR kiribati OR korea OR kosovo OR kyrgyzstan OR kirghizia OR "Kyrgyz Republic" OR kirghiz OR kirgizstan OR "Lao PDR" OR laos OR lebanon OR lesotho OR basutoland OR liberia OR libya OR macedonia OR madagascar OR "Malagasy Republic" OR malaysia OR malaya OR malay OR sabah OR sarawak OR malawi OR mali OR "Marshall Islands" OR mauritania OR mauritius OR "Agalega Islands" OR mexico OR micronesia OR "Middle East" OR moldova OR moldovia OR moldovian OR mongolia OR montenegro OR morocco OR ifni OR mozambique OR myanmar OR myanma OR burma OR namibia OR nepal OR "Netherlands Antilles" OR "New Caledonia" OR nicaragua OR niger OR nigeria OR pakistan OR palau OR palestine OR panama OR paraguay OR peru OR philippines OR philipines OR phillipines OR phillippines OR "Puerto Ric*" OR romania OR rumania OR

roumania OR rwanda OR ruanda OR "Saint Lucia" OR "St Lucia" OR "Saint Vincent" OR "St Vincent" OR grenadines OR samoa OR "Samoan Islands" OR "Navigator Island" OR "Navigator Islands" OR "Sao Tome" OR senegal OR serbia OR montenegro OR seychelles OR "Sierra Leone" OR "Sri Lanka" OR "Solomon Islands" OR somalia OR "South Africa" OR sudan OR suriname OR surinam OR swaziland OR syria OR tajikistan OR tadzhikistan OR tadjikistan OR tadzhik OR tanzania OR thailand OR togo OR togolese republic OR tonga OR tunisia OR turkey OR turkmenistan OR turkmen OR uganda OR ukraine OR uzbekistan OR uzbek OR vanuatu OR "New Hebrides" OR venezuela OR vietnam OR "Viet Nam" OR "West Bank" OR yemen OR yugoslavia OR zambia OR zimbabwe))) OR (TITLE-ABS-KEY ("Developing Countries" OR africa OR asia OR caribbean OR "West Indies" OR "South America" OR "Latin America" OR "Central America" OR ((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income" OR underserved OR "under served" OR deprived OR poor*) W/1 (countr* OR nation* OR population* OR world)))) OR (TITLE-ABS-KEY (((developing OR "less* developed" OR "under developed" OR underdeveloped OR "middle income" OR "low* income") W/1 (economy OR economies)) OR (low* W/1 (gdp OR gnp OR "gross domestic" OR "gross national")) OR (low W/3 middle W/3 countr*))) OR (TITLE-ABS-KEY (((lmic OR lmics OR "third world" OR "lami countr*")) OR "transitional countr*"))) AND ((TITLE ("Food security" OR poverty OR "household* income*") OR ABS ("Food security" OR poverty OR "household* income*") OR TITLE ((increas* OR improv* OR lower* OR decreas* OR diminish* OR reduc* OR loss OR declin* OR slump OR dwindl* OR curtail* OR restrict* OR shrink* OR fall) W/3 (income* OR revenue* OR yield* OR productivity)) OR ABS ((increas* OR improv* OR lower* OR decreas* OR diminish* OR reduc* OR loss OR declin* OR slump OR dwindl* OR curtail* OR restrict* OR shrink* OR fall) W/3 (income* OR revenue* OR yield* OR productivity)) OR TITLE ("market power" OR net-return* OR "net return*" OR outcome* OR effect* OR impact) OR ABS ("market power" OR net-return* OR "net return*" OR outcome* OR effect* OR impact)) AND (TITLE (contract* OR "nucleus estate*" OR cooperative* OR "producer* association*") OR ABS (contract* OR "nucleus estate*" OR cooperative* OR "producer* association*") OR TITLE (embedded W/3 service*) OR ABS (embedded W/3 service*) OR TITLE ((pre-harvest) W/2 (agreement* OR sales)) OR ABS ((pre-harvest) W/2 (agreement* OR sales)) OR TITLE ("value chain*" OR farm-firm* OR outgrow*) OR ABS ("value chain*" OR farm-firm* OR outgrow*) OR TITLE ((vertical) W/3 (integration OR coordination OR linkage*)) OR ABS ((vertical) W/3 (integration OR coordination OR linkage*))) AND ((TITLE (farm* OR smallhold* OR "small hold*" OR small-hold*) OR ABS (farm* OR smallhold* OR "small hold*" OR small-hold*) OR TITLE ((small-scale OR "small scale"

) W/3 (producer*)) OR ABS ((small-scale OR "small scale") W/3 (producer*))) OR (TITLE (agricultur* OR outgrower* OR "small farmer*" OR "small grower*") OR ABS (agricultur* OR outgrower* OR "small farmer*" OR "small grower*")) OR (TITLE ((vegetable* OR fruit OR livestock OR dairy OR milk OR beef OR poultry OR pig* OR flower* OR cereal OR tea OR soybean* OR rice OR coffee OR potato* OR sugarcane OR mushroom* OR maize OR millet OR pepper* OR crop OR crops) W/3 (produc* OR grow*)) OR ABS ((vegetable* OR fruit OR livestock OR dairy OR milk OR beef OR poultry OR pig* OR flower* OR cereal OR tea OR soybean* OR rice OR coffee OR potato* OR sugarcane OR mushroom* OR maize OR millet OR pepper* OR crop OR crops) W/3 (produc* OR grow*))) OR (TITLE (floriculture) OR ABS (floriculture)))))) **- 973 hits**

CAB Abstracts search – Searched 29th September 2015

- Database: CAB Abstracts

19 ("Food security" or poverty or "household* income*" or ((lower* or decreas* or diminish* or reduc* or loss or declin* or slump or dwindl* or curtail* or restrict* or shrink* or fall) adj3 (income* or revenue* or yield* or productivity)) or "market power" or net-return* or "net return*" or outcome* or effect* or impact).ti,ab,sh. (2095228)

21 (Contract* or "nucleus estate*" or cooperative* or "producer* association*" or (embedded adj3 service*) or (pre-harvest adj2 (agreement* or sales)) or "value chain*" or farm-firm* or outgrow* or (vertical adj3 (integration or coordination or linkage*))).ti,ab,sh. (52881)

23 (Farm* or smallhold* or "small hold*" or small-hold* or ((small-scale or "small scale") adj3 producer*) or agricultur* or outgrower* or "small farmer*" or "small grower*" or ((vegetable* or fruit or livestock or dairy or milk or beef or poultry or pig* or flower* or cereal or tea or soybean* or rice or coffee or potato* or sugarcane or mushroom* or maize or millet or pepper* or crop or crops) adj3 (produc* or grow*)) or floriculture).ti,ab,sh. (943886)

24 (Afghanistan or Angola or Albania or "American Samoa" or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Belarus or Belize or Benin or Bolivia or Bosnia or Herzegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burkina Fasso or Burundi or Urundi or Cambodia or Cameroon or Cameroons or Cameron or Camerons or Central African Republic or Chad or China or Colombia or Comoros or Comoro Islands or Comores or Congo or Costa Rica or Cuba or Zaire or Cote d'Ivoire or Ivory Coast or Djibouti or Dominica* or East Timor or East Timur or Timor Leste or Ecuador or Egypt or United Arab Republic or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gambia or Gaza or Georgia Republic or Georgian Republic or Ghana or Grenada or Guatemala or Guinea or Guiana or Guyana or Haiti or Honduras or Hungary or India or Indonesia or Iran or Iraq or Kazakhstan or Kenya or Kiribati or

Korea or Kosovo or Kyrgyzstan or Kirghizia or Kyrgyz Republic or Kirghiz or Kirgizstan or Lao PDR or Laos or Lebanon or Lesotho or Liberia or Libya or Macedonia or Madagascar or Malagasy Republic or Malawi or Malaysia or Maldives or Marshall Islands or Mali or Mauritania or Mauritius or Agalega Islands or Mexico or Micronesia or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or Nicaragua or Niger or Nigeria or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philipines or Phillipines or Phillippines or Romania or Rwanda or Ruanda or Samoa or Samoan Islands or Sao Tome or Senegal or Serbia or Seychelles or Sierra Leone or Sri Lanka or Solomon Islands or Somalia or South Africa or St Lucia or St Vincent or Grenadines or Sudan or Suriname or Swaziland or Syria or Tajikistan or Tadzhikistan or Tadjikistan or Tadzhik or Tanzania or Thailand or Tonga or Togo or Togolese Republic or Tunisia or Turkey or Turkmenistan or Tuvalu or Uganda or Ukraine or Uzbekistan or Uzbek or Vanuatu or Venezuela or New Hebrides or Vietnam or Viet Nam or West Bank or Yemen or Zambia or Zimbabwe).hw,ti,ab,cp. (2004287)

25 ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world)).ti,ab. (44493)

26 ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab. (712)

27 (low* adj (gdp or gnp or gross domestic or gross national)).ti,ab. (43)

28 (low adj3 middle adj3 countr*).ti,ab. (1881)

29 (lmic or lmics or third world or lami countr*).ti,ab. (2264)

30 transitional countr*.ti,ab. (78)

31 exp developing countries/ (1347769)

32 or/24-31 (2068603)

33 22 and 23 and 32 (9796)

34 19 and 22 and 23 and 32 (3496)

****

Annex 3 - Coding tool

	ID	Question	Description	Coding
FIRST CODING BLOCK - EPPI REVIEWER CODING RELEVANT FOR ALL QUANTITATIVE STUDIES INCLUDED AFTER FULL-TEXT SCREENING				
General	ID	Unique study identification #	EPPI	EPPI Internal ID
	ABSTRACT	Abstract	Abstract (EPPI)	Text
	PUB DATE	Publication date	Year of publication of study (EPPI)	Year
	PUB TYPE	Publication type	Publication type (EPPI)	1= Peer-reviewed journal 2= Book chapter/book 3= Unpublished report 4 = Conference proceedings 5 = Other
	RELEVANCE	Relevance	Relevance	1 = Relevance 2 = Methodology 3 = Relevance and methodology
Study design	STUDY DESIGN	Design type	Randomised experiment makes comparisons with a control group measured in the same time period as the treatment group, with randomised assignment of the contract to each group; Quasi-experiment means assignment to intervention within treatment group is random, but treatment and intervention groups are non-equivalent groups. Observables are used to match treatment and comparison groups, e.g. PSM, instrumental variables, covariate matching, regression discontinuity designs, etc. Non-experiment do not use a control group nor multiple measurements in time.	1= Randomised experiment 2= Quasi-experiment 3= Non-experiment
			If quasi-experiment: what design is used (e.g. PSM, DID, IV)	Specify (text)

	ID	Question	Description	Coding
	COMPARISON	Comparison group of farmers	Farmers in the comparison group are selling independent (no group selling),	1= Yes 2= No 8= Not clear 9= N/A
			Farmers in the comparison group have no contract with any firm	1= Yes 2= No 8= Not clear 9= N/A
			Farmers in the comparison group sell all their produce in the traditional or spot market	1= Yes 2= No 8= Not clear 9= N/A
	MEASUREMENT FREQ	Measurement frequency	How often were farmers in the intervention and control group interviewed and compared?	1= Comparison at single point in time 2= Comparison at two points in time 3= Comparison at more than two points in time
	INTERVENTION PERIOD	Length of follow-up	How much time has passed between the start of the contract farming arrangement and the measurement of effects?	Specify (textt)
	STUDY TYPE	Study type	Does the study refers to an empirical instance of contract farming with a comparisons group or other credible counterfactual design to measure effectiveness?	1 = Yes = Quantitative effectiveness study useful for meta-analysis 2 = No = Quantitative effectiveness study not useful for meta-analysis
	REL-INFO	Rel-info	Reasons for exclusion from meta-analysis	Specify (text)
END FIRST CODING BLOCK				

	ID	Question	Description	Coding
SECOND CODING BLOCK - EPPI REVIEWERONLY FOR STUDIES INCLUDED FOR META-ANALYSIS				
Intervention	CONTRACT PARTY	Contract party	The type of organization the farmer signed the contract with	1= Private firm 2= Farmers cooperative 3= NGO 4= Other 8= Not clear
	CONTR PARTY OTH	Contract party if other	The type of organization the farmer signed the contract with	Open answer
	CONSTITUYENTS	Is study in its effectiveness analysis considering smallholders?	YES = Reporting needs to comply with at least two out of four criteria 1. Limited size of farm (reported as compared to other farms in the sector) 2. Mostly dependent on family labour 3. Subsistence farming or mix of subsistence and market-oriented farming 4. Reportedly limited resources in terms of land, technical and technological support, and/or capital for maintenance and investment.	1= Yes 2= No 8= Not clear
	DEF-SMALLH	Criteria used	1. Limited size of farm (reported as compared to other farms in the sector) 2. Mostly dependent on family labour 3. Subsistence farming or mix of subsistence and market-oriented farming 4. Reportedly limited resources in terms of land, technical and technological support, and/or capital for maintenance and investment.	List 1, 2, 3....
		Definition of smallholder	Definition of smallholder used in the study	Specify (text)
	FORMAL	Written or verbal	Whether the contract is written or verbal	1= Written 2= Verbal 8= Not clear
	PREPLANTING	Pre-planting	Contracted before production begins?	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
	RESOURCE PROVISIONING	Seed provisioning	Any seed or plant provisioning included in the contract conditions	1= Yes 2= No 8= Not clear
		Agro-chemicals provisioning	Any agro-chemicals provisioning included in the contract conditions	1= Yes 2= No 8= Not clear
		Credit provisioning	Any credit provisioning in money included in the contract conditions (do not include seeds, agro-chemicals, or extension here)	1= Yes 2= No 8= Not clear
		Other resource/service provided	Other implements, tools, infrastructure, etc.	Specify (text)
	TRAINING	Extension provisioning	Any extension provided to contract farmers	1= Yes 2= No 8= Not clear
		Extension provider	Type of organisation providing the extension	1= Private firm 2= Farmers cooperative 3= NGO 4= Other 8= Not clear 9= N/A
	PROCESS OR PRDUCT CONDITIONS	Production conditions	Process conditions included in the contract, regarding input use, crop varietal, health and safety regulations, labour regulations, etc.	1= Yes 2= No 8= Not clear 9= N/A
		Marketing conditions	Product conditions included in the contract, regarding quality, timing of delivery, food safety (residues), traceability, etc.	1= Yes 2= No 8= Not clear 9= N/A
		Environmental conditions	Environmental conditions included in the contract, regarding on-farm biodiversity, water and waste management, pollution, etc.	1= Yes 2= No 8= Not clear 9= N/A
		Other conditions	Other proocess or product conditions mentioned in the study	Specify (text)

	ID	Question	Description	Coding
	PAYMENT	Price determination and payment system	Agreed price determination and payment system?	1= Yes 2= No 8= Not clear 9= N/A
	DISPUTE	Dispute resolution	Is a dispute resolution mechanism mentioned?	1= Yes 2= No 8= Not clear 9= N/A
	RIGHTS	Rights of firm	Is there a non-transferable right and legal title to the crop for the firm?	1= Yes 2= No 8= Not clear 9= N/A
Mapping empirical instances	SECTOR	Type of crop	Type of crop under contract	1 = Annual crops perishable 2 = Annual crops non-perishable 3 = Perennial crops 4 = Animal husbandry 5 = Multiple/mixed
	CROPS	Contracted crops	The crops contracted	Crop 1, crop 2, etc
	REGION	Region	Region of empirical instance of contract farming arrangement	1 = Africa 2 = Latin America 3 = Asia 4 = Europe 5 = Oceania 6 = North America
	COUNTRY	Country	Countries of empirical instance of contract farming arrangements in study	Country 1, country 2, etc
	TYPE	Type of contractual arrangement	Type of contract farming (multiple categories allowed)	1 = Contract farming only 2 = Certification included 3 = Forward sales included 4 = Collective marketing included 5 = Risk insurance included 6 = Access to support services included

	ID	Question	Description	Coding
Outcomes	ULTIMATE OUTCOMES	Ultimate outcomes	Household income	1 = Reported 2 = Not reported 8 = Unclear
			Household food security	1 = Reported 2 = Not reported 8 = Unclear
			Market power	1 = Reported 2 = Not reported 8 = Unclear
			Livelihood strategies of contracting farmers	1 = Reported 2 = Not reported 8 = Unclear
			Positive spill-over effects on non-contracting farmers (income, knowledge, ..)	1 = Reported 2 = Not reported 8 = Unclear
			Increased interest from (formerly) non participating farmers in becoming a contracted farmer	1 = Reported 2 = Not reported 8 = Unclear
			Negative spill-over effects on non-contracting farmers (pollution, exclusion, ..)	1 = Reported 2 = Not reported 8 = Unclear
			Other	Specify
	INTERMEDIATE OUTCOMES	Intermediate outcomes	Agricultural practices	1 = Reported 2 = Not reported 8 = Unclear
			Yields or other productivity measure	1 = Important 2 = Not important 8 = Unclear importance
			Crop revenue, profit, or other variable on crop income	1 = Important 2 = Not important 8 = Unclear importance
			Financial risk exposure	1 = Important 2 = Not important 8 = Unclear importance

	ID	Question	Description	Coding
			Health and food safety	1 = Reported 2 = Not reported 8 = Unclear
			Other	Specify
IMPACT	DEVELOPMENT IMPACT	Agricultural sector development	Poverty alleviation	1 = Reported 2 = Not reported 8 = Unclear
			Economic growth	1 = Reported 2 = Not reported 8 = Unclear
			National food security	1 = Reported 2 = Not reported 8 = Unclear
			Employment	1 = Reported 2 = Not reported 8 = Unclear
			Other	Specify
END SECOND CODING BLOCK				

	ID	Question	Description	Coding
THIRD CODING BLOCK - EPPI REVIEWERRISK OF BIAS TOOL FOR OUTCOMES COVERED IN THE META-ANALYSIS				
RISK OF BIAS (Hombrados and Waddington, 2014)	1. Was bias in the mechanism of assignment prevented?	RCT	Yes = a random component in the sequence generation process is described (e.g. referring to a random number table); and if the unit of allocation was at group level (geographical/ social/ institutional unit) and allocation was performed on all units at the start of the study, or if the unit of allocation was by beneficiary or group and there was some form of centralised allocation mechanism such as an on-site computer system; and if the unit of allocation is based on a sufficiently large sample size to equate groups on average. UNCLEAR= the paper does not provide details on the randomisation process, or uses a quasi-randomization process for which it is not clear has generated allocations equivalent to true randomisation. NO = the sample size is not sufficient or any failure in the allocation mechanism could affect the randomisation process.	1= Yes 2= No 8= Not clear
		Regression Discontinuity Designs	Yes = allocation is made based on a pre-determined discontinuity on a continuous variable (regression discontinuity design) and blinded to participants or, if not blinded, individuals reasonably cannot affect the assignment variable in response to knowledge of the participation decision rule; and the sample size immediately at both sides of the cut-off point is sufficiently large to equate groups on average. UNCLEAR = the assignment variable is either non-blinded or it is unclear whether participants can affect it in response to knowledge of the allocation mechanism.	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
			NO = if: · the sample size is not sufficient or there is evidence that participants altered the assignment variable prior to assignment	
		non-randomised programme placement and self-selection (including matching strategies, excluding IV)	Yes = Participants and non-participants are either matched based on all relevant characteristics explaining participation and outcomes, or all relevant characteristics are accounted for. UNCLEAR = it is not clear whether all relevant characteristics (only relevant time varying characteristics in the case of panel data regressions) are controlled. NO = relevant characteristics are omitted from the analysis.	1= Yes 2= No 8= Not clear
		Instrumental variable estimation	YES = An appropriate instrumental variable is used which is exogenously generated: e.g. due to a 'natural' experiment or random allocation. UNCLEAR = the exogeneity of the instrument is unclear (both externally as well as why the variable should not enter by itself in the outcome equation). NO = otherwise	1= Yes 2= No 8= Not clear
	2. Was bias due to group equivalence prevented?	RCT	YES = baseline characteristics of the study and control/-comparisons are reported and overall similar based on t-test or ANOVA for equality of means across groups, or covariate differences are controlled using multivariate analysis; and the attrition rates are sufficiently low and similar in treatment and control, or the study assesses that loss to follow up units are random draws from the sample; and problems with cross-overs and drop outs are dealt with using intention-to-treat analysis or in the case of	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
			drop outs, by assessing whether the drop outs are random draws from the population; and, for cluster-assignment, authors control for external cluster-level factors that might confound the impact of the programme through multivariate analysis. UNCLEAR = insufficient details are provided on covariate differences or methods of adjustment; or insufficient details are provided on cluster controls. NO = otherwise	
		Regression Discontinuity Designs	YES = the interval for selection of treatment and control group is reasonably small, or authors have weighted the matches on their distance to the cut-off point, and the mean of the covariates of the individuals immediately at both sides of the cut-off point (selected sample of participants and non-participants) are overall not statistically different based on t-test or ANOVA for equality of means, or significant differences have been controlled in multivariate analysis; and, for cluster-assignment, authors control for external cluster-level factors that might confound the impact of the programme (eg weather, infrastructure, community fixed effects, etc) through multivariate analysis. UNCLEAR = there are covariate differences across individuals at both sides of the discontinuity which have not been controlled for using multivariate analysis, or if insufficient details are provided on controls, or if insufficient details are provided on cluster controls. NO = otherwise	1= Yes 2= No 8= Not clear
		statistical matching	YES = the authors use a difference-in-differences (or	1= Yes

	ID	Question	Description	Coding
		studies including propensity scores (PSM) and covariate matching	fixed effects) multivariate estimation method; the authors control for a comprehensive set of time-varying characteristics; · and the attrition rate is sufficiently low and similar in treatment and control, or the study assesses that drop-outs are random draws from the sample (e.g. by examining correlation with determinants of outcomes, in both treatment and comparison groups); and, for cluster-assignment, authors control for external cluster-level factors that might confound the impact of the programme (eg weather, infrastructure, community fixed effects, etc) through multivariate analysis. UNCLEAR = insufficient details are provided, ·or if insufficient details are provided on cluster controls. No = otherwise	2= No 8= Not clear
		regression-based studies using cross sectional data (excluding IV)	YES = matching is either on baseline characteristics or time-invariant characteristics which cannot be affected by participation in the programme; and the variables used to match are relevant (e.g. demographic and socio-economic factors) to explain both participation and the outcome (so that there can be no evident differences across groups in variables that might explain outcomes) In addition, for PSM Rosenbaum's test suggests the results are not sensitive to the existence of hidden bias. and, with the exception of Kernel matching, the means of the individual covariates are equated for treatment and comparison groups after matching; and, for cluster-assignment, authors control for external cluster-level factors that might confound the impact of the programme (eg weather, infrastructure, community fixed effects, etc) through multivariate or any appropriate analysis.	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
			UNCLEAR = relevant variables are not included in the matching equation, or if matching is based on characteristics collected at endline, or if insufficient details are provided on cluster controls. NO = otherwise	
		instrumental variables approaches	YES = the instrumenting equation is significant at the level of F≥10 (or if an F test is not reported, the authors report and assess whether the R-squared (goodness of fit) of the participation equation is sufficient for appropriate identification); the identifying instruments are individually significant (p≤0.01); for Heckman models, the identifiers are reported and significant (p≤0.05);where at least two instruments are used, the authors report on an over-identifying test (p≤0.05 is required to reject the null hypothesis); and none of the covariate controls can be affected by participation and the study convincingly assesses qualitatively why the instrument only affects the outcome via participation and, for cluster-assignment, authors particularly control for external cluster-level factors that might confound the impact of the programme (eg weather, infrastructure, community fixed effects, etc) through multivariate analysis. UNCLEAR = relevant confounders are controlled but appropriate statistical tests are not reported or exogeneity13 of the instrument is not convincing, or if insufficient details are provided on cluster controls (see category f) below). NO = otherwise	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
	3. Is bias due to Hawthorne and John Henry effects prevented?	Was the process of being observed causing motivation bias?	YES = a) For data collected in the context of a particular intervention trial (randomised or non-randomised assignment), the authors state explicitly that the process of monitoring the intervention and outcome measurement is blinded, or argue convincingly why it is not likely that being monitored in ways that could affect the performance of participants in treatment and comparison groups in different ways. b) The study is based on data collected in the context of a survey, and not associated with a particular intervention trial, or data are collected in the context of a retrospective (ex post) evaluation. UNCLEAR = it is not clear whether the authors use an appropriate method to prevent Hawthorne and John Henry Effects (e.g. blinding of outcomes and, or enumerators, other methods to ensure consistent monitoring across groups). No = otherwise	1= Yes 2= No 8= Not clear
	4. Are SPILL-OVER effects properly addressed?	Was the study adequately protected against spill-over bias?	YES = the intervention is unlikely to spill-over to comparisons (e.g. participants and non-participants are geographically and/or socially separated from one another and general equilibrium effects are unlikely) UNCLEAR = spill-overs are not addressed clearly. NO = allocation was at individual or household level and there are likely spill-overs within households and communities which are not controlled for in the analysis; or if allocation at cluster level and there are likely spill-overs to comparison clusters.	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
	5. Selective outcome reporting:	Was the study free from outcome reporting bias?	YES = there is no evidence that outcomes were selectively reported (e.g. all relevant outcomes in the methods section are reported in the results section). NO = some important outcomes are subsequently omitted from the results or the significance and magnitude of important outcomes was not assessed. UNCLEAR = otherwise	1= Yes 2= No 8= Not clear
	6. Selective analysis reporting:	Was the study free from analysis reporting bias?	YES = authors use 'common' methods15 of estimation and the study does not suggest the existence of biased exploratory research methods. For PSM and covariate matching, score "YES" if: - Where over 10% of participants fail to be matched, sensitivity analysis is used to re-estimate results using different matching methods (Kernel Matching techniques). - For matching with replacement, no single observation in the control group is matched with a large number of observations in the treatment group. For IV (including Heckman) models, score "YES" if: - the authors test and report the results of a Hausman test for exogeneity ($p \leq 0.05$ is required to reject the null hypothesis of exogeneity). - the coefficient of the selectivity correction term (Rho) is significantly different from zero ($P<0.05$) (Heckman approach). For studies using multivariate regression analysis, score "YES" if: - authors conduct appropriate specification tests (e.g. reporting results of multicollinearity test, testing robustness of results to the inclusion of additional variables, etc). UNCLEAR = unreported	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
			NO = authors use uncommon or less rigorous estimation methods such as failure to conduct multivariate analysis for outcomes equations where it is has not been established that covariates are balanced.	
	7. Other bias	Was the study free from other sources of bias?	YES = the reported results do not suggest any other sources of bias. UNCLEAR = other important threats to validity may be present NO = it is clear that these threats to validity are present and not controlled for.	1= Yes 2= No 8= Not clear
	8. Statistical conclusion validity	For studies using parametric regression methods such as OLS: Are variables properly scales?	YES = the authors test and fail to reject the null of homoscedasticity (e.g. through a Breusch-Pagan test for heteroscedasticity (p>0.05)) and test for the assumed error distribution (e.g. Kolmogorov-Smirnov test for non-normality (p>0.05)) or if the test suggests the existence of heterogeneity or non-normality, the study corrects for them (e.g. use of log transformation in the dependent variable). UNCLEAR = the results of any test are not reported. NO = otherwise	1= Yes 2= No 8= Not clear
		Is the study sufficiently powered to detect the effects?	YES = the sample size is enough to detect a relevant significant effect. UNCLEAR = it is not clear whether the sample size is sufficiently large to detect medium or large significant effects. NO = the sample size is not sufficiently large to detect medium or large significant effects.	1= Yes 2= No 8= Not clear

	ID	Question	Description	Coding
		For clustered studies: Does prevent unit of analysis errors?	YES = the analysis is carried out at the relevant unit of treatment assignment, or the study accounts for lack of independence between observations within assignment clusters. UNCLEAR = the study does not report enough information on the unit of treatment assignment. NO = the analysis is carried out at a different unit than the assignment.	1= Yes 2= No 8= Not clear
Meta-analysis	METAN (per outcome area - per proxy-indicator with the least risk of bias)	Meta-analysis	Sample size treatment; Outcome measure in treatment; SD of outcome measure in treatment	STATA- metan; EPPI-meta-analysis
			Sample size comparison; Outcome measure in comparison; SD of outcome measure in comparison	
			Type of effect: Average treatment effect (ATE), Average treatment on the treated (ATT), Local average treatment effect (LATE), Intention to treat effect (ITT)	1 = ATE 2 = ATT 3 = LATE 4 = ITT 8 = Other
END THIRD CODING BLOCK				

	ID	Question	Description	Coding
FOURTH CODING BLOCKDATA EXTRACTION FOR EACH OUTCOME AREA IN EXCEL DATA-BASE				
META-ANALYSIS	OUTCOME DESCIRPTION	OUTCOME DEF	Does the study give a precise definition of outcome?	1=Yes 2=No
		DEFINITION	What definition of outcome is given?	Open answer
		METRIC	What metric is used in the calculation of the outcome variable used in estimation?	
		S SIZE METRIC	Sample size unit of measurement	1= Farmers 2= Households 3= Groups (e.g. village) 4= Other 5= Sample size not stated
		OTHER - METRIC	If other, state what	Open answer
	SAMPLE SIZE	S SIZE TREAT	Initial sample size treatment group	#
		S SIZE CONTR	Initial sample size control group	#
		ATTRIT TREAT	Number of drop-outs	#
		ATTRIT CONTR	Number of drop-outs	#
		ATTRIT COMMENT	Are assessments made of whether drop-outs are unlikely to be randomly selected from sample? If so, state covariates reported	Open answer
		SAMPLE SIZE	Total number of beneficiaries in study (after attrition)	#
		EST SAMPLE	Where sample size not provided in study, give estimated number	#
	EFFECT	ITT	Does the study use ITT analysis?	1=Yes 2=No 9= N/A
		TOT	Does the study use TOT analysis?	1=Yes 2=No 9= N/A
		ES TYPE	Type of effect estimate calculated for meta-analysis	1=Risk/rate 2=Standardised mean difference 3=Other
		ES - OTHER	If other, state what	Open answer
	OUTCOME VALUES	BASELINE T	Baseline outcome - Treatment	#
		BASELINE C	Baseline outcome - Comparison	#
		ENDLINE T	Endline outcome - Treatment	#
		ENDLINE C	Endline outcome - Comparison	#
	SAMPLE SIZE			#
	ES CALC	Calculated ES	Calculated effect-size	#
	CI L	CI lower	95% Confidence interval lower bound reported (X.XX)	#

	ID	Question	Description	Coding
	CI U	CI upper	95% Confidence interval upper bound reported (X.XX)	#
	ES SE	Standard error of RR	Standard error of estimate	#
	CORRECT SE	Standard errors	Are standard errors appropriately adjusted for?	1=Yes 2=No 8=Not clear 9= N/A
	SE COMMENTS	Comments on S.E. calculation	Comments on standard error calculation, as reported in paper or as calculated from data in paper	Open answer
END FOURTH CODING BLOCK				

	ID	Question	Description	Coding
FIFTH CODING BLOCK - ALTAS.TIQUALITATIVE SYNTHESIS ON EMPIRICAL INSTANCES COVERED IN META-ANALYSIS(NOTE: ALSO OTHER STUDIES WITH COMPLEMENTARY INFORMATION ON THE SAME EMPIRICAL INSTANCES ARE CODED				
BARRIERS AND ENABLERS	STARTING CONDITIONS	Company attributes	Organisation of production	Text/verbatim
			Nationality	Text/verbatim
			Ownership	Text/verbatim
			Oher	Specify (text)
		Product attributes	Perishability	Text/verbatim
			Production technology	Text/verbatim
			Economies of scale	Text/verbatim
			Other	Specify (text)
	ENABLING ENVIRONMENT	Geographic conditions	Climate	Text/verbatim
			Soil type	Text/verbatim
			Access to water	Text/verbatim
			Access to markets	Text/verbatim
			Other	Specify (text)
		Market conditions	Markets and competition	Text/verbatim
			Market regulatiosn	Text/verbatim
			History of economic shocks	Text/verbatim
			Other	Specify (text)
		Business environment	Stakeholder partnerships	Text/verbatim
			Business support services	Text/verbatim
			Presence of financial services	Text/verbatim
			Third party quaity verification	Text/verbatim
			Reliable legal system	Text/verbatim
			Other	Specify (text)
	PRODUCT CHARACTERISTICS	Crop characteristics	Yields (quality & quantity)	Text/verbatim
			Threshold investments	Text/verbatim
			Profitability	Text/verbatim
			Self-consumption	Text/verbatim
			Other	Specify (text)

	ID	Question	Description	Coding
	FARMER CHARACTERISTICS	Smallholder characteristics	Human capital	Text/verbatim
			Social capital	Text/verbatim
			Physical capital	Text/verbatim
			Financial capital	Text/verbatim
			Natural capital	Text/verbatim
			Other	Specify (text)
	SOCIAL CHARACTERISTICS	Social structure	Social relations of production (e.g. class, differentiation, farming structure)	Text/verbatim
			Cultural heritage (e.g. caste, identity, common pool resources)	Text/verbatim
			Other	Specify (text)
INCENTIVE STRUCTURE	FIRM INCENTIVES	Incentive mentioned in the impact study	Reliable quality and quantity	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Economies of scale in procurement	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Control of production and process and crop attributes	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Off-loading production risk	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Reputation (e.e. licence to operate)	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Intelectual property rights	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Contract management costs	1 = Relevant 2 = Not relevant 8 = Unclear relevance

	ID	Question	Description	Coding
			Cost of non-compliance	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Other	Specify (text)
		Contract implementation issues	Contract conditions used, refined or breached	Specidy(text_
			Other	Specify (text)
	FARMER INCENTIVES	Farmer incentive structure	Expected higher income	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Technical upgrading	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Access to markets	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Price risks	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Production risks	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Social risks	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Financial risks	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Increased market risks	1 = Relevant 2 = Not relevant 8 = Unclear relevance
			Other	Specify (text)
		Contract implementation issues	Contract conditions used, refined or breached	Specify (text)
			Other	Specify)text)
END FIFTH CODING BLOCK				